LITERATURE REVIEW: Do Adult Females with High Levels of Eating Pathology have difficulties Recognising Emotions in Faces? A Systematic Review

EMPIRICAL PAPER: Emotion Recognition and Set Shifting in Women with Anorexia Nervosa

Submitted by **Royston Hall**, to the University of Exeter as a thesis for the degree of **Doctor of Clinical Psychology**, May 3rd 2018

Signature: 

Author's Declaration

Royston Hall completed the literature review independently. For the empirical paper, the emotion recognition task used was provided by Dr Ian Penton-Voak of the University of Bristol who was involved in the development and training of the use of the present task. Data collection, analysis and write-up of the empirical paper were completed independently by Royston Hall.

Table of Contents

List of Tables

Literature Review

Empirical Paper

List of Figures

Literature Review

Empirical Paper

SCHOOL OF PSYCHOLOGY

DOCTORATE IN CLINICAL PSYCHOLOGY

LITERATURE REVIEW

Do Adult Females with High Levels of Eating Pathology have difficulties Recognising Emotions in Faces? A Systematic Review

Trainee Name	**Royston Hall**
Primary Research Supervisor (50%)	**Dr Nick Moberly** Senior Lecturer, School of Psychology, University of Exeter
Secondary Research Supervisor (50%)	**Dr Ian Frampton** Clinical Psychologist, Eating Disorders Service, Cornwall NHS Foundation Trust
Target Journal	European Eating Disorders Review.
Word Count	5965 words (excluding abstract, table of contents, tables and figures, references, footnotes, appendices)

Submitted in partial fulfilment of requirements for the Doctorate Degree in Clinical Psychology, University of Exeter Signature:

Abstract

Background*:* Individuals with high eating pathology may also demonstrate difficulties with certain forms of social cognition including emotion recognition of faces (ER), however, this relationship is not well understood.

Objective*:* The purpose of this review was to determine whether eating pathology is associated with ER difficulties in adult females.

Method*:* Case control studies comparing ER performance between groups of high eating pathology (e.g. a diagnosed eating disorder) and low eating pathology were selected from multidisciplinary databases published prior to February 2018. The systematic literature search yielded 396 articles which were screened in line with PECOS criteria. Of 62 full text screened papers, 20 met all criteria and were synthesised into the present review.

Results*:* Results confirmed an overall trend for reduced ER performance in high eating pathology groups, though the specific nature of ER differences was variable across studies. Differences in outcome were not due to confounding factors but instead inconsistency in ER tasks employed.

Conclusions*:* The review supports the position that there are differences in ER among individuals with high eating pathology compared to low eating pathology, however, the exact nature of these differences are not fully understood. Further research is still required as heterogeneity of ER measures limits the current interpretation of the literature.

Keywords: *eating pathology, emotion recognition, social cognition, systematic review*

Introduction

This review explores the relationship between emotion recognition (ER) ability and eating pathology in adult women. While themes of ER ability in specific eating disorders (ED) have been explored in the past (Caglar-Navali et al., 2014; Oldershaw et al., 2011), a full evaluation of the literature across all forms of eating pathology has not been conducted. This review aims to consider the strengths and limitations of the present evidence while critically summarising the methodology of the studies in this field.

Social Cognition and Eating Pathology

Eating disorders (EDs) are characterised by disturbed or inappropriate eating behaviours often underpinned by cognitive distortions relating to body image and weight (American Psychiatric Association [APA], 2013). Many factors have been indicated in the development and maintenance of EDs such as anorexia nervosa (AN) and bulimia nervosa (BN) including neurobiological vulnerabilities (Kaye, Wierenga, Bailer, Simmons, & Bischoff-Grethe, 2013), sociocultural factors (Witztum, Latzer, & Stein, 2008) and psychological factors (Dellava et al., 2010). One unifying construct that is pertinent to biological, psychological and social theories of eating pathology is social cognition (Oldershaw et al., 2011). Social cognition (SC) is defined as our ability to recognise, understand and regulate our social and emotional experience (Adolphs, 1999). SC deficits are associated with poor psychological well-being creating vulnerability for the development of EDs (Adenzato, Todisco, & Ardito, 2012; Tchanturia et al., 2012) as well as mood difficulties more generally (Cardi et al., 2018; Ladegaard, Videbech, Lysaker, & Larsen, 2015; Weightman, Air, & Baune, 2014). Indeed, SC ability may have the most

significant impact on recovery rates in EDs (Speranza, Loas, Wallierr, & Scorco, 2007). As such, these skills have become central to several cognitive and neuropsychological models of EDs (Cooper, Todd, & Wells, 2009; Schmidt & Treasure, 2006).

SC entails a variety of complex and interconnected cognitive skills (see Oschner, 2008), of which facial emotion recognition (ER) remains the most well researched yet contentious measure in eating disorder research. The link between ER and eating pathology may not be immediately apparent, however, one key concept that may link these two seemingly distinct constructs is emotion regulation (Brockmeyer et al., 2014). Effective emotion regulation skills are dependent on our ability to accurately recognise the emotional state of others and recognise and manage our own emotional state in response. Poor understanding of emotional experience of the self and others may lead to emotional avoidance (Treasure & Schmidt, 2013) or inhibition (Bekker & Spoor, 2008) strategies. For example, food restriction in eating disorders may serve to distance the self from the reception of unwanted or unmanageable negative emotions (Corstorphine, Mountford, Tomlinson, Waller, & Meyer, 2007). Thus, difficulties in ER may be a contributor towards emotion regulation difficulties that perpetuate eating pathology (Haynos & Fruzzetti, 2011). Indeed, ER ability has been found to be associated with reduced emotion regulation skills in an eating disorder population (Harrison et al., 2009). Furthermore, several studies have demonstrated that ER ability is significantly reduced in women with EDs (Jänsch, Harmer, & Cooper, 2009; Kucharska-Pietura, Nikolaou, Masiak, & Treasure, 2004). However, successful replication of this finding has been limited (Cardi et al., 2015; Kessler et al., 2006) and there is also inconsistency into the specific nature of ER difficulties reported in

ED. Indeed, studies have reported specific difficulties in recognising happy (Jones, Harmer, Cowen, & Cooper, 2008), fearful (Kucharska-Pietura et al., 2004), sad (Castro, Davies, Hale, Surguladze, & Tchanturia, 2004) and disgusted faces (Lulé et al., 2014) in ED populations.

Challenges to Interpreting Emotion Recognition Literature

Variability in co-morbid factors may account for some inconsistency in ER reported among ED populations. ER ability may be associated with depression (Bourke, Douglas, & Porter, 2010; Dalili, Penton-Voak, Harmer, & Munafo, 2015) and anxiety (Demenescu, Kortekaas, Den Boer, & Aleman, 2010), both of which are highly co-morbid among ED populations (Blinder, Cumella, & Sanathara, 2006; Swinbourne et al., 2012). Another comorbidity relevant to ER is alexithymia: difficulty in identifying and describing one's own emotional state to others and difficulty distinguishing between internal feelings and bodily sensations (Taylor, 1994). Alexithymia is commonly reported in ED populations (Nowakowski, McFarlane, & Cassin, 2013) and may contribute towards poor ER performance (Lane, Sechrest, Riedel, Shapiro, & Kaszniak, 2000).

As well as comorbidity, illness factors also need to be considered in ER research. For example, it is unclear if ER difficulties are a trait of ED (Harrison, Tchanturia, & Treasure, 2010) or are a state brought on by the illness itself that will resolve upon recovery (Oldershaw, Hambrook, Tchanturia, Treasure, & Schmidt, 2010). It is possible that variation in the severity of illness (e.g., BMI) of participants in ER studies may impact

upon cognitive functions due to neurophysiological changes associated with starvation (McCormick et al., 2008). Similarly, ER performance may be related to changes associated with age and neurological maturation (Frith & Frith, 2003; Horning, Cornwell, & Davis, 2012; Isaacowitz et al., 2007), particularly in the frontal lobes (Amodio & Frith, 2006), as well as the hormonal effects of puberty (Lawrence, Campbell, & Skuse, 2015).

Finally, variability between studies may be a result of the heterogeneity of measures used to capture ER performance. Most experimental tasks use a forced recognition response to emotional face stimuli based on static images of the six universal facial expressions of emotion (Ekman & Friesen, 1971). As well as inherent ecological validity limitations, adjustments of various task parameters in this simple design may significantly alter the outcomes measured. For example, experimental factors including the length of stimuli presentation, the method of response and range of emotions measured may alter the way participants perform.

Thus, the present review aims to comprehensibly synthesise studies evaluating emotion recognition ability of women with disordered eating including AN, BN and Eating Disorder Not Otherwise Specified (EDNOS). To review whether ER difficulties are related to a state of illness or represent a vulnerability trait, non-clinical studies will be included where high eating pathology groups are well defined and represent an ‘at-risk’ group that has not previously received a diagnosis of an eating disorder. Therefore, the current systematic review aims to answer the question “do adult females with high levels of eating pathology have difficulties recognising emotions in faces?”

Methods

Eligibility Criteria

Studies identified through the search strategy described below were screened by title and abstract. In line with PRISMA-P recommendations (Moher et al., 2015), the remaining full texts were included if they met specific conditions based on PECOS criteria (Methley et al., 2014) set out in Table 1. Full text of six randomly selected papers were assessed using the same eligibility criteria by a second rater. No disagreement was found on the eligibility of the identified papers. Due to the under-representation of men in eating disorder research, only studies with an exclusively female population are included. Similarly, studies with adult participants (age 18 or above) were included to minimise the impact of neurodevelopment on the synthesis of results. In clinical studies, a professional confirmed DSM-5 diagnosis of eating disorder in high eating pathology groups. In non-clinical studies, self-report measures such as the eating disorders examination questionnaire or equivalent (EDE-Q; Berg, Peterson, Frazier, & Crow, 2012; Fairburn & Beglin, 1994) determined the presence of high eating pathology. All studies were required to evaluate ER in high eating pathology groups against a control group with lower reported pathology as determined by lack of ED diagnosis or lower EDE-Q self-report measures. As this comparison requires a case-control study design, only such designs were included in the synthesis. Grey literature was searched through reference lists of identified papers and through google scholar to reduce the impact of publication bias (McAuley, Tugwell, &

Moher, 2000). However, a systematic search of the grey literature was not possible due to resource limitations.

Table 1.

Inclusion and exclusion criteria based on PECOS guidelines (Moher et al., 2015)

Parameter	**Inclusion Criteria**	**Exclusion Criteria**
Population	Participants aged 18 or older. Female.	
Exposure	Study assesses eating disorder diagnosis or pathological eating behaviour (e.g., restricting or bingeing) regardless of severity or clinical significance.	
Comparison	Healthy controls without eating disorder diagnosis and/or individuals with lower levels of disordered eating / eating pathology on continuous scale.	No low eating pathology control condition to compare. No correlation examined with continuous measures of disordered eating / eating pathology.
Outcomes	Study examines ER in facial expression. Validated and well described experimental measure for determining emotion recognition ability.	Study solely examines other forms of emotional functioning (e.g., emotion regulation or emotional expression).
Study Design	Case control studies	Case reports. Editorials and opinion pieces. Systematic reviews.

Information Sources

Four databases were searched: PubMed, PsycINFO, PsycARTICLES and Web of Science Core Collection®. Also, a mnual search of the reference list and the cited list of all included studies in Google Scholar was conducted. Results of the search were taken from the “title” fields as abstract searching identified a significant number of irrelevant studies. The search included all articles published from the start date of each database up until February 2018.

Search Terms

The current review was conducted in line with PRISMA-P guidelines (Moher et al., 2015). Search terms were determined from keywords of seminal publications and related systematic reviews in similar fields (Caglar-Nazali et al., 2014; Oldershaw et al., 2011). Search terms can be found in Table 2.

Table 2.

Search terms used in systematic literature review of databases

Construct	Individual search terms
Section 1 Participants Eating Pathology	(“anorexi*” “bulimi*” “eating disorder*” “eating pathology” “disordered eating” “bing*”)
Section 2 Emotion	(“emot*” “affect” “happy” “happiness” “surpris*” “fear*” “anger” “angry” “expression*” “sad*” “disgust*” “expression*” “face*” “facial” “social”)
Section 3 Recognition	(“Recog*” “identif*” “social judgment*” “judgement*” “processing” “theory of mind” “cogniti”)
Search:	Section 1 AND section 2 AND section 3

Data extraction

Data extracted from the studies was organised and synthesised within an excel sheet with information placed in columns of appropriate headings (e.g. sample size, experimental task details and so on). Excel allowed for easy organisation and management of data to observe trends within the findings of the studies reviewed.

Study evaluation tool

All studies extracted by the systematic search were subjected to an evaluation of their quality using the Quality Assessment Tool for Quantitative Studies (QATQS, Thomas, Ciliska, Dobbins, & Micucci, 2004) from the Effective Public Health Project (Armijo-Olivo, Stiles, Hagen, Biondo, & Cummings, 2012; Appendix A). In addition to eligibility, the quality of six studies selected at random from the studies included in the review were assessed by a second independent rater using the same tool. No disagreement was identified in any domains of the studies analysed.

Results

394 records were identified through database searches described above (See figure 1). A further two studies were identified through reference lists, bringing the total to 396. 210 duplicates were removed, and a further 124 papers did not meet inclusion criteria as screened by title and abstract. Of the remaining 62 papers, 42 did not meet eligibility criteria as determined by PECOS leaving a final 20 papers included in the review (see table 3). The populations reported within the studies included AN (*n* = 13), BN (*n* = 8),

EDNOS (*n* = 2) and non-clinical (*n* = 4). Apart from one doctoral thesis (Marcon, 2017), all studies included in the review were published in peer-reviewed journals.

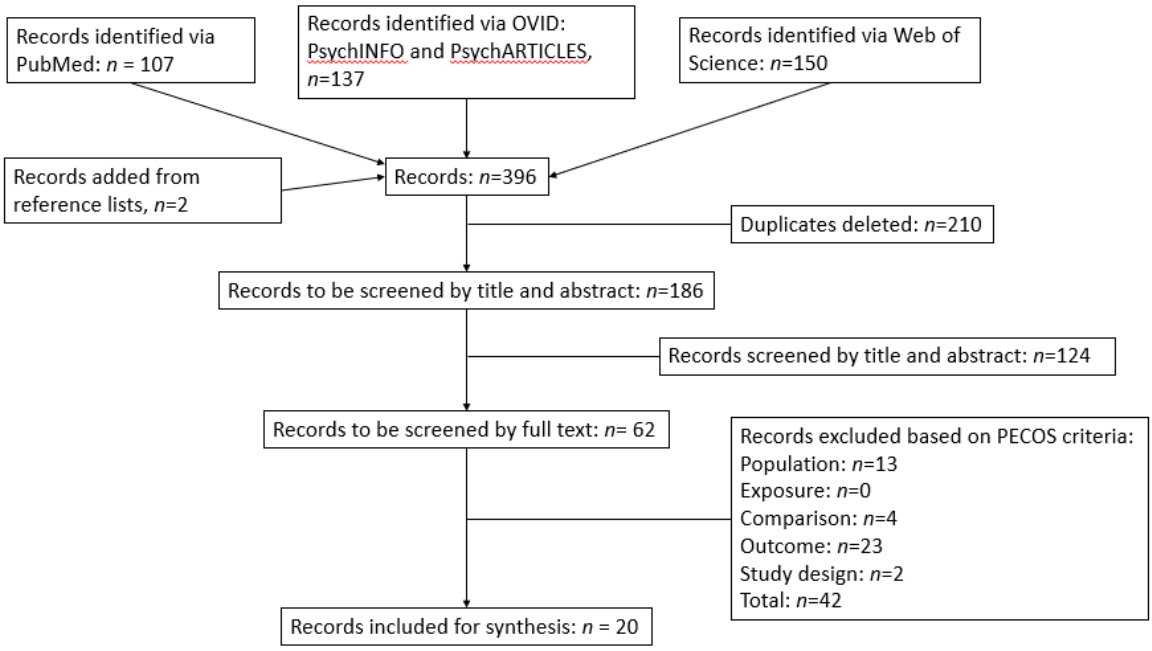

Figure 1. Results of literature search strategy and eligibility screening.

Table 3.

Summary of all reviewed papers

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
1	Kucharska-Nikolaou, Vailis Masiak, & Treasure, 2002	AN = 30 Mean age = 20.2 YEARS Mean duration of illness = 33.9 months No psychopathology measures reported. DSM-IV criteria used – HC = 30 Mean age = 25.2 years Age ranges not reported. RecruitmentClinical: Inpatient ED unit Control: University staff and students	Emotion Recognition Experiment (ERE): 36 emotional face photos. Forced choice response: Anger Sadness Surprise Fear Happy Disgust Contempt Shame Interest	Global differences in ER: AN<HC	Effect sizes not reported	A = Strong B = Moderate C = Strong D = Moderate E = Weak F = Moderate **GLOBAL = Moderate**	**Strengths** complex emotional expressions included beyond the universal six (e.g. contempt and shame). Good validity and test-retest reliability of task. **Limitations** Long presentation times for each stimulus (10 seconds) reducing ecological validity of task. Use of medication not reported. Samples not matched by age or education.
2	Russell, Schmidt, Doherty, Young, & Tchanturia, 2009	AN = 22 Mean age = 26.7 years Age range = 20-35 Mean duration of illness = 9.5 years No psychopathology measures reported. DSM-4 criteria used – HC = 22 Mean age = 30.3 years Age range = 20-46 Recruitment Clinical: Inpatient ED unit Control: Local advertisement	Reading the Mind in the Eyes (RME): 25 photographs of the eye-region displayed. Forced choice of four emotion or thought words.	Global difference in ER: AN<HC AN performed worse on items with female eyes	*d* = 1.37 (very large) *d* = .99 (large)	A = Strong B = Moderate C = Moderate D = Moderate E = Strong F = Moderate **GLOBAL = Strong**	**Strengths** Analysed impact of gender of face stimuli on recognition performance. Used a well validated and reliable measure of ER. Samples matched by IQ. **Limitations** Relatively small sample size. Clinical and non-clinical groups were not matched by age. or education Use of medication not reported.

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
3	Pollatos, Herbert, Schandry, & Gramann, 2008	AN = 15 Mean age = 23.9 years Mean duration of illness = 3.7 years SCID-I measures confirmed eating pathology of clinical group. DSM-5 criteria used HC = 15 Mean age = 22.4 years Age ranges not reported <u>Recruitment</u> Clinical: Patient self-help groups Control: Local advertisement	Bespoke task: 40 faces from the Karolinska Directed Emotional Faces database presented in random order. Forced choice response: Anger Sadness Fear Happy Disgust Neutral	Global differences in ER: AN<HC Neutral, sad, disgusted: AN<HC Neutral face performance associated with alexithymia	η_p^2 = .19 (small) No effect size reported. *r* = .61 (large)	A = Moderate B = Moderate C = Moderate D = Moderate E = Strong F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Samples matched by age, education and medication use (excluded if on medication) between groups. Explored recognition alongside neuroimaging data. **<u>Limitations</u>** Very small sample size. Somewhat self-selecting sample as participants were recruited from a self-help group. Did not measure surprise recognition. Reliability and validity of measure not reported.
4	Jänsch , Harmer, & Cooper, 2009	AN = 28 Mean age = 27.1 years EAT-26 measure confirmed eating pathology of clinical group. DSM-5 criteria used HC =28 Mean age = 28.2 years Age ranges and length of illness not reported <u>Recruitment</u> Clinical: Inpatient ED unit Control: Not stated	Facial Expression Recognition Task (FERT): emotions each displayed at 10 levels of intensity (0-100%). Forced choice response: Anger Sadness Surprise Fear Happy Disgust Neutral	Global differences in ER: AN<HC Specific differences in ER: Neutral, sad, disgusted: AN<HC Depression scores influenced performance on ER task	No effect size reported.	A = Strong B = Moderate C = Moderate D = Moderate E = Moderate F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Used morphed facial stimuli increasing ecological validity. Well-matched samples by age and education. **<u>Limitations</u>** Sample varied on medication use and was too small for comparison of medication status. Reliability and validity of measure not reported.

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
5	Harrison, Sullivan, Tchanturia, & Treasure, 2009	AN = 20 Mean age = 26.3 years Mean duration of illness = 7.2 years EDDS measure confirmed eating pathology of clinical group. DSM-IV criteria used HC = 20 Mean age = 28.3 years Age ranges not reported <u>Recruitment</u> Clinical: Inpatient and outpatient ED unit Control: Local advertisement	Reading the Mind in the Eyes (RME): 25 photographs of the eye-region displayed. Forced choice of four emotion or thought words	Global difference in ER: AN<HC ER performance correlates with self-reported emotion regulation skills	*d* = 1.2 (very large) *r* = -.35 (medium)	A = Strong B = Moderate C = Strong D = Moderate E = Strong F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Well matched sample by age and IQ. Well validated and reliable measure of ER. **<u>Limitations</u>** Relatively small sample size. Medication use not reported.
6	Brewer, Cook, Cardi, Treasure, & Bird, 2015	AN = 21BN = 2 Mean age = 23.4 years EDE-Q measure confirmed eating pathology of clinical group. DSM-5 criteria used HC = 21 Mean age = 25.7 years Age range and length of illness not reported <u>Recruitment</u> Not stated	Bespoke: Emotional faces presented with stepwise white noise obscuring image. Yes/no forced choice response. Emotion recognition thresholds of white noise recorded on each emotion: Anger Sadness Surprise Fear Happy Disgust Pain Neutral	No significant difference in ER performance between AN and HC. ER performance correlated with alexithymia.	η_p^2= .00 (very small) *r* = -.34 (medium)	A = Strong B = Moderate C = Strong D = Moderate E = Moderate F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Novel task, no ceiling effects due to adjustment of task based on response. Samples matched based on age. IQ and alexithymia. **<u>Limitations</u>** All male face stimuli. Mixed sample of ED. Medication use not reported. Relatively small sample size.

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
7	Gramaglia et al., 2016	AN = 39 Mean age = 30.6 years SCID-I measure confirmed eating pathology of clinical group. HC = 48 Mean age = 32 years Age range and length of illness not reported RecruitmentClinical: Inpatient and outpatient ED unit Control: University staff and students	Facial Emotion Identification Test (FEIT): 55 emotion faces, forced choice response: Anger Sadness Surprise Fear Happy Disgust Neutral	Specific differences in ER: fear: AN<HC disgust: AN>HC	*d*=0.68 (large) *d*=0.58 (large)	A = Strong B = Moderate C = Moderate D = Moderate E = Moderate F = Moderate **GLOBAL = Strong**	**Strengths** Included both static and dynamic facial displays of emotion. Samples matched by age. **Limitations** Long presentation time of facial stimulus (15 second) reducing ecological validity of task. Low reliability of emotion recognition tasks. Did not control for medication use. Samples not matched by education.
8	Dapelo, Surguladze, Morris, & Tchanturia, 2016	AN = 35 Mean age = 27.5 yearsMean duration of illness = 10.5 SCID-I measure confirmed eating pathology of clinical group. DSM-IV criteria used HC = 42 Mean age = 26.9 years Age ranges not reported Recruitment Clinical: b-eat website advertisement Control: Local advertisement	Bespoke: Pairwise blended faces at 5 proportions (e.g. 90:10, 70:30, 50:50 etc) presented. Forced choice response: Anger Sadness Fear Happy Disgust	Specific differences in ER: Disgust at 90% proportion: AN<HC, Anger at 90% proportion: AN>HC,	*r*=-0.39 (medium) *r*=0.36 (medium)	A = Moderate B = Moderate C = Strong D = Moderate E = Moderate F = Moderate **GLOBAL = Strong**	**Strengths** Novel task with blended facial expressions with greater ecological validity. Emphasis on response bias as well as correct responses. Role of medication use on ER explored. Samples matched by age and education. **Limitations** Skewed data (ceiling effect). Non-parametric analysis. Reliability and validity of ER task not reported though stimuli taken from a validated source.

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
9	Kessler, Schwarze, Filipic, & Traue, 2006	AN = 48 Mean age = 22.9 years Mean hours of previous outpatient treatment = 67.4 BN = 31Mean hours of previous outpatient treatment = 93.8 Mean age = 25.5 years Eating pathology measure not reported. DSM-IV criteria used HC = 78 Mean age = 22.8 years Age ranges not reported RecruitmentClinical: Inpatient ED unit Control: University and nursing student	Facially Expressed Emotion Labelling (FEEL): presentation of static emotion images takes from validated databases. Forced choice response: Anger Sadness Surprise Fear Happy Disgust	Specific differences in ER: Surprise: AN+BN<HC	No effect size reported.	A = Strong B = Moderate C = Moderate D = Moderate E = Strong F = Moderate **GLOBAL = Strong**	**Strengths** Large sample size, well matched groups. High reliability reported for ER task. Samples matched by age. **Limitations** Mixed AN/BN sample without sub-group analysis. Sample not matched by education. Medication use not reported.
10	Harrison, Sullivan, Tchanturia, Treasure, 2010	AN= 50 Mean age = 26.7 years Mean years of illness = 9.2BN = 50 Mean age = 27.5 years Mean years of illness = 8.4 EDE-Q measure confirmed eating pathology of clinical group. DSM-IV criteria used HC = 90 Mean age = 28.5 Age ranges not reports RecruitmentClinical: Inpatient ED units Control: University staff and students	Reading the Mind in the Eyes (RME): 25 photographs of the eye-region displayed. Forced choice of four emotion or thought words.	Global difference in ER: AN<HC+BN, ER correlated with severity of illness as reported on EDE-Q, but not on BMI,	*r* = .21 (medium) *r* = .37 (medium) *r* = .35 (medium)	A = Moderate B = Moderate C = Strong D = Moderate E = Strong F = Moderate **GLOBAL = Strong**	**Strengths** Well matched groups by age, education and IQ. Large sample size. Good reliability and validity of ER task reported. Role of medication use on ER explored. **Limitations** Individual emotions not explored within task. Self-selecting sample. Unlimited stimulus exposure time reducing ecological validity

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
11	Rothschild, Eviatar, Shamia, & Gur, 2011	AN = 9BN = 20 Mean age of clinical sample combined = 23.3 years EAT-26 measure confirmed eating pathology of clinical group. DSM-IV criteria used HC = 27 Mean age = 23.1 years Age ranges and length of illness not reported <u>Recruitment</u> Clinical: Inpatient ED unit Control: University students	Bespoke (unnamed): Judgement as to whether two faces share the same or different emotive expressions. 40 trials: Happy Anger	No global difference in emotion recognition by group	*d* = .03 (small)	A = Strong B = Moderate C = Moderate D = Moderate E = Weak F = Moderate **GLOBAL = Moderate**	**<u>Strengths</u>** Robust measures of social functioning. Samples matched by age. **<u>Limitations</u>** ER task had no time limit reducing ecological validity. Only anger and happiness assessed. Samples not matched by IQ or education. Medication use not reported.
12	Dapelo, Surguladze, Morris, & Tchanturia, 2017	AN = 35 Mean age = 27.5 years Mean length of illness = 10.5 years BN = 26 Mean age = 26.4 years Mean length of illness = 7.7 years SCID-I and EDE-Q measures confirmed eating pathology of clinical group. DSM-5 criteria used HC = 42 Mean age = 26.9 years Age ranges not report <u>Recruitment</u> Clinical: Inpatient ED units and b-eat Control: University students	Bespoke: Pairwise blended faces at 5 proportions (e.g. 90:10, 70:30, 50:50 etc) presented. Forced choice response: Anger Sadness Fear Happy Disgust	Specific differences in ER at 90% proportion: Disgust: AN+BN<HC Anger: AN+BN>HC	*r* = -.39 (medium) *r* = .36 (medium)	A = Strong B = Moderate C = Moderate D = Moderate E = Moderate F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Comparably large sample size. Novel task using blended facial expressions with greater ecological validity. Emphasis on response bias as well as correct responses. Samples matched by age. **<u>Weaknesses</u>** Skewed data (ceiling effect). Non-parametric analysis. Education, IQ and medication use not reported. Reliability and validity of ER task not reported.

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
13	Medina-Pradas, Navarro, Alvarez-Moya, Grau, & Obiols, 2012	AN = 44 Mean age = 26.8 years Mean length of illness = 9.9 years BN = 30 Mean age = 26.8 years Mean length of illness = 10.3 years EDNOS = 39 Mean age = 26 years Mean length of illness = 7.3 years SCID-I measure confirmed eating pathology of clinical group. DSM-IV criteria used HC = 20 Mean age = 26 years Age ranges not reported <u>Recruitment</u> Clinical: Inpatient ED unit Control: Not specified	Reading the Mind in the Eyes (RME): 25 photographs of the eye-region displayed. Forced choice of four emotion or thought words.	Global difference in ER but not in AN: BN<HC EDNOS<HC AN=HC	*d* = .64 (medium) *d* = .77 (medium) *d* = 1.09 (Large) *d* = .19 (very small)	A = Strong B = Moderate C = Strong D = Moderate E = Strong F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Included three main eating disorder sub-types. Samples matched by age. Good reliability and validity of ER task reported. Comparably large sample size. **<u>Limitations</u>** Individual emotions not explored within task. Unlimited presentation time reduces ecological validity. Samples not matched by education and medication use not reported.

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
14	Legenbauer, Vocks, & Rüddel , 2008	BN = 20 Mean age = 22.6 years EDE measure confirmed eating pathology of clinical group. DSM-IV criteria used HC = 20Mean age = 23.9 years Age ranges and length of illness not reported. <u>Recruitment</u> Clinical: Inpatient ED units Control: University staff and students	Bespoke: taken from the Japanese and Caucasian Facial Expression of Emotion (JACFEE) and neutral faces (JACNeuF). Forced choice response: Anger Sadness Surprise Fear Happy Disgust Contempt Neutral	Specific differences in ER: Surprise: BN<HC	*d* = .86 (large)	A = Strong B = Moderate C = Strong D = Moderate E = Strong F = Moderate **GLOBAL =Strong**	**<u>Strengths</u>** Employed a control task and neutral faces to explore recognition bias. Samples matched by age. Good reliability and validity of ER task reported. **<u>Limitations</u>** Small sample size. Bias data not collected on main ER task. Long stimulus display time may reduce ecological validity. Education and IQ not reported.
15	Kühnpast , Gramann, Pollatos, 2012	BN = 16 Mean age = 24.6 years Mean duration of illness = 8.1 years EDI measure confirmed eating pathology of clinical group. DSM-IV criteria used HC = 13 Mean age = 25.4 years Age ranges not reported <u>Recruitment</u> Clinical: Inpatient and outpatient ED units Control: University students	Bespoke: faces selected from the Karolinska Directed Emotional Faces image battery. Forces choice response: Anger Fear Happy Neutral	Global difference in ER: BN<HC Specific differences in ER: Angry: BN<HC	η_p^2 = .61 (large) No effect size reported.	A = Strong B = Moderate C = Strong D = Moderate E = Weak F =Moderate **GLOBAL = Moderate**	**<u>Strengths</u>** Novel electrophysiological data presented in combination with ER performance. Samples matched by age, education and medication use. **<u>Limitations</u>** Very small sample size. Part of a larger neuroimaging study, participants may have been fatigued at time of participation. Small selection of emotions analysed Reliability and validity of ER task not reported.

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
16	Kenyon et al., 2012	BN = 48 Mean age = 28 years Mean duration of illness = 10.7 years EDNOS = 34 Mean age = 27.6 years Mean duration of illness = 9.9 years EDE-Q measure confirmed eating pathology of clinical group. DSM-IV criteria used HC = 57 Mean age = 24 Age ranges not reported <u>Recruitment</u> Clinical: Inpatient and outpatient ED unit Control: University staff and students	Reading the Mind in the Eyes (RME): 25 photographs of the eye-region displayed. Forced choice of four emotion or thought words.	Specific differences in ER: Negative emotions: BN>HC	η_p^2 = .01 (small)	A = Strong B = Moderate C = Strong D = Moderate E = Strong F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Compared positive and negative emotions in analysis. Good reliability and validity of ER task reported. Role of medication use on ER explored. Comparatively large sample size. **<u>Limitations</u>** Individual emotions not explored within task. Unlimited presentation time reduces ecological validity. Non-parametric analysis. Samples not matched by age or IQ. Education status not reported.
17	Jones, Harmer, Cowen, Cooper, 2008	HC = 52 High EAT-26 (Score 16 or greater) = 29 Mean age = 23.4 years Low EAT-26 (Score of 3 or less) = 23 Mean age = 27 years Age ranges not reported <u>Recruitment</u> University staff and students	Facial Expression Recognition Task (FERT): emotions each displayed at 10 levels of intensity (0-100%). Forced choice response: Anger Sadness Surprise Fear Happy Disgust Neutral	Specific differences in ER: Happy: High EDI<low EDI, Neutral: High EDI<low EDI,	*r* = .57 (medium) *r* = .74 (medium)	A = Moderate B = Moderate C = Strong D = Strong E = Strong F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Well matched groups by age and education. Participants screened if on medication. Morphed stimuli with greater ecological validity. Good reliability and validity of ER task reported. **<u>Limitations</u>** Self-selected sample. Groups split by self-reported eating pathology

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
18	Ridout, Wallis, Autwal, Sellis, 2012	HC = 80 High EDI (Median split: <17) = 40 Mean age = 26.4 years Low EDI (Median split >18) = 40 Mean age = 22.8 years Age ranges not reported <u>Recruitment</u> University staff and students – self-help centre for eating concerns	Bespoke: adaption of Facial Expression of Emotion: Stimuli and Tests (FEEST): emotions each displayed at one of four intensities. Forced choice response: Anger Sadness Fear Happy Disgust	Global difference in ER: High EDI<low EDI High EDI groups misidentified more fearful faces as anger	η_p^2 = .1 (medium) No effect size reported.	A = Moderate B = Moderate C = Moderate D = Strong E = Moderate F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Looked at specific pattern of recognition errors. Varied intensity of emotion expression to increase ecological validity. Samples matched by education. **<u>Limitations</u>** Undergraduate population. Samples not matched by age. Excluded surprise. Self-selection of participants. Exclusion criteria of participants not reported. Reliability and validity of ER task not reported.
19	Sharpe, Wallis, & Ridout, 2016	HC = 35 High EDI =18 Mean age = 21.2 years Low EDI = 17 Mean age = 21.7 years Age ranges not reported <u>Recruitment</u> Not stated	Bespoke: Brief task with 14 faces. Force choice response: Anger Sadness Surprise Fear Happy Disgust Neutral	Global difference in ER: High EDI<low EDI Specific differences in ER: Fear: High EDI<low EDI Neutral: High EDI<low EDI	η_p^2 = .11 (medium) *d* = .71 (medium) *d* = .97 (large)	A = Moderate B = Moderate C = Moderate D = Strong E = Moderate F = Moderate **GLOBAL = Strong**	**<u>Strengths</u>** Novel eye-tracking approach presented alongside ER data. Sample matched by age. **<u>Limitations</u>** Small sample size and small number of face stimuli trials used. Questionnaires completed before experimental tasks which could prime negative mood. Delineation of sample groups not described in detail. Educational history or IQ not reported and exclusion criteria not reported. Reliability and validity of ER task not reported.

No.	Author	Sample	ER Measurement	Key Findings	Effect Sizes	Quality Assessment	Strengths and Weaknesses
20	Marcon, 2017	HC = 77 Mean age of total sample = 21.8 High EAT-26 = 38 Low EAT-26 = 39 Mean age for High and Low Eat-26 groups not reported Age range = 18-63 <u>Recruitment</u> University students	Bespoke: taken from the (JACFEE). Each emotion presented at one of four intensities. Not forced choice. Images were: Anger Sadness Surprise Fear Happy Disgust	No significant global difference in emotion recognition by group.	No effect size reported.	A = Moderate B = Moderate C = Weak D = Strong E = Moderate F = Moderate **GLOBAL = Moderate**	**<u>Strengths</u>** ER task not forced-choice response, stimuli morphed increasing ecological validity. **<u>Limitations</u>** No demographic information presented. Unlimited presentation time reduces ecological validity. Reliability and validity of ER task not reported. Delineation of sample groups not described in detail.

Note: AN = anorexia Nervosa, BN = bulimia nervosa, EAT-26, Eating Attitudes Test, EDDS = Eating Disorder Diagnostic Scale, EDE-Q Eating Disorder Examination – Questionnaire, EDI = Eating Disorder Inventory, EDNOS = eating disorder not otherwise specified, HC = healthy control; SCID-I Structure Clinical Interview for DSM, QATQS=quality assessment tool for quantitative studies: A=selection bias, B=study design, C=confounders, D=blinding, E=data collection method, F=withdrawals and dropouts.

Critical Summary

Quality of Papers

The twenty studies were published between the years of 2002 and 2017. Of the sixteen studies that used clinical samples, fourteen recruited participants directly from ED services. One study advertised solely through an eating disorder charity website (8) and one study did not describe their recruitment procedure (6). Control group participants were predominantly university undergraduates which may limit the generalisability of these studies (1, 11, 14, 15, 17, 18, 20). Several studies also recruited control groups through community advertisements (2, 3, 4, 5, 7, 8, 10, 12, 16) while three studies did not detail the recruitment process for control groups (6, 13, 19). Total sample sizes varied from 29 to 190 (mean 76.5, median 58). Seven of the studies reviewed may be considered underpowered due to samples smaller than would be required to detect a large effect (d < .8 with power set at .80) in an analysis comparing two groups (N < 52; 2, 3, 5, 6, 14, 15, 19). These papers should be interpreted carefully due to the increased risk of type II errors. Similarly, the reduced reliability of such studies may mean a higher rate of type I errors in the literature base associated with publication bias towards significant results. Indeed, only one of the six studies with a small sample size did not find a difference in ER between groups (6), although this study did have a noticeably different experimental task that did not require forced labelling of emotional states.

The critique of studies was guided by the Quality Assessment Tool for Quantitative Studies (QATQS) (Thomas et al., 2004) which provides an overall quality

rating of weak, moderate or strong (see appendix A) as well as ratings in a series of domains relevant to quantitative research. Most studies received strong ratings in each domain; however, there were some notable exceptions. In the domain of selection bias, four studies using non-clinical populations could only be rated as moderate as groups were defined by a single parameter (self-reported eating pathology) (17, 18. 19, 20). Similarly, five papers received medium ratings for selection bias as there was a degree of self-selection for the clinical sample (3, 8, 10, 17, 18). Concerning the management of confounders, nine studies received moderate ratings, and one study was rated as weak (20). The moderate and weak ratings were awarded due to differences between groups that were not evaluated or discussed further in the analysis. Finally, nine studies received a moderate rating for data collection methods while three studies were given a weak rating due to inadequate or unreported reliability and validity of tasks used (1, 11, 15). The overall rating was determined by the number of weak ratings across all domains. Four studies received an overall rating of moderate (1, 11, 15, 20), while the remaining sixteen were rated as strong. All reviewed studies were based on a similar case-review design and consequently factors such as blinding, withdrawal and drop-out could not be evaluated, limiting the variability of quality ratings reported here.

Global findings of ER studies

Across all twenty studies, ten identified a global reduction in the number of correctly identified emotional expressions in high eating pathology groups (1, 2, 3, 4, 5, 10, 13, 15, 18, 19). Effect sizes for global differences were medium (10, 13, 18, 19) to large (2, 5, 15) with only one study reporting a small effect size (3). A further seven papers identified specific ER differences between groups, explored in more detail below

(7, 8, 9, 12, 14, 16, 17). Studies that found specific differences predominantly reported medium (8, 12, 17) to large (7, 14) effect sizes, however, one small effect size was also reported (16). Three studies in this review found no relationships between emotion recognition and eating pathology (6, 11, 20). Three studies did not report sufficient data to calculate effect sizes (1, 4, 9).

Of the ED populations examined, AN was the most represented. Seven out of thirteen studies reported global ER difficulties in the AN group such that they recognised significantly fewer emotions in faces than controls (1, 2, 3, 4, 5, 10, 13). Effect sizes reported were small (3), medium (10, 13) and large (2, 5). Two studies did not report effect sizes (1, 4). Four of the thirteen AN studies found specific ER differences between AN and controls (discussed below) and only two studies failed to reject the null hypothesis of no difference between AN and controls on ER tasks (6, 11). Both studies that found no difference in ER performance in the AN group were the only studies that did not require forced choice labelling of emotions. These particular tasks required participants to provide yes or no responses to stimuli. Evaluating stimuli by several criteria (e.g. picking from multiple emotion labels) is likely to be more challenging than evaluating by one criterion (happy or not happy), potentially contributing towards the lack of effect observed in the latter studies.

Eight studies also compared ER performance of individuals with BN, of which three studies reported a global difference in the number of correctly identified emotions in BN (10, 13,15). Effect sizes of these studies varied from medium (10, 13) to large (15). Three other studies identified specific ER differences (12, 14, 16), while only one study failed to identify a relationship between BN and ER tasks (11). Two studies

included a sample of EDNOS participants in their analysis, one found a global difference in ER between groups (13), while another found comparable performance with controls (16).

Finally, four studies measuring non-clinical eating pathology were included. In all four studies, the sample was split by high or low self-reported eating pathology on measures such as the EDI (Garner, 1991). One study used a median split (18), two studies employed a tertile split (17, 19) and one study did not specify the stratification of groups, though groups were orthogonal on EDI measures (20). Two studies demonstrated a global reduction in ER performance in high EDI groups compared to low EDI (18, 19), while one study found specific ER differences between groups (17). Effect sizes for these studies were all medium. Only one study did not find any differences between high and low EDI groups (20). This study was the only task using a non-clinical population that did not have a presentation limit for stimuli, potentially reducing the ecological validity of the present task and making it substantially easier.

Specific findings of ER Differences

The specific differences in emotion recognition ability between high and low eating pathology groups varied significantly with trends only observable for disgust and anger. For disgust, three out of seven studies demonstrated reduced recognition in AN (3, 8, 12). One of these studies also demonstrated reduced disgust recognition in BN as well as AN (12). However, of these seven studies, one study did not follow this trend and found that the AN group were significantly better at recognising disgust than controls (7).

Only one in four studies comparing BN and one in four studies comparing non-clinical eating populations reported specific difficulties recognising anger in the ED group (15, 18). Interestingly, four of seven studies comparing AN populations identified the opposite trend where anger was better differentiated by the high eating pathology group (5, 8, 10, 12).

Reduced surprise recognition was only reported in one of five AN studies (9) and in the sole study to explore this emotion in BN (14). Contrary to this, high EDI scores in one of two non-clinical populations better-differentiated surprise than the low EDI group (17). Recognition of fearful faces was found to be reduced in one of seven studies exploring AN (7), and two of four studies comparing non-clinical eating pathology (18, 19). Only one of seven studies reported a trend for reduced recognition of sadness in AN (3) and in non-clinical populations one of four studies reported reduced recognition of happiness in the high EDI group (17). Finally, reduced recognition of neutral faces was reported in one of four AN population studies (3) as well as high EDI groups in two non-clinical population studies (17, 19).

Taken together, there is a consensus of an overall reduced ER ability in high eating pathology compared to low eating pathology. However, except for a trend for poor disgust and greater anger recognition in AN, there remains considerable variation in the specific differences in ER and eating pathology with no other trends apparent.

Confounding Variables

Age of participants was recorded and matched across groups in all but four studies (1, 2, 16, 18) and there was a small range of mean ages across samples (21.7

to 31.9). Four studies matched IQ between groups (2, 5, 6, 10) while two studies had significantly higher IQ in the non-clinical group (11, 16). The association of IQ on ER ability was inconsistent; while some studies found these variables to be independent (2, 6, 11), two studies found direct correlations between IQ and ER performance in ED populations (10, 16) and one study did not report the impact of IQ on ER performance (5). Both studies that found this relationship used the National Adult Reading Test (NART; Nelson & Willison, 1991) as a measure of premorbid IQ in addition to an ER task requiring a sophisticated vocabulary for emotional words (Reading the Mind in the Eyes [RME]; Baron-Cohen Jolliffe, Mortimore, & Robinson, 1997). It is possible that the association observed between these two tasks is mediated by language ability. h Eight studies matched groups based on years of education while five studies were unable to match groups by education (1, 2, 7, 9, 13). Seven studies did not record or report years of education within their sample (5, 6, 12, 14, 16, 19). No studies reported a relationship between education and ER ability. Overall there was no robust relationship between IQ and education variables and the strength of ER differences between groups.

BMI measurements were reported in 14 studies. In three BN studies, BMI was matched across groups (14, 15, 16), suggesting that differences found in these studies are not be attributable to the effects of low body weight. While BMI was significantly lower in most clinical groups, nine studies explored associations between BMI and ER and demonstrated that weight status did not influence performance between groups (2, 3, 6, 8, 10, 13, 14, 16, 17). This finding is perhaps not surprising as the severity of eating pathology and BMI are likely to share variance. Four studies demonstrated that length of illness did not predict ER performance (3, 8, 10, 13). Of four studies examining

medication, three did not demonstrate an impact of medication on ER (8, 16, 17), while the fourth found an anecdotal effect of medication on ER performance (4).

Sixteen studies recorded measures of depression through validated questionnaire measures and found higher rates of depression in high eating pathology groups. All studies that measured depression controlled for this construct in their analyses. Most papers (*n* = 15) found no relationship between depression symptoms and ER ability; however, one study identified a specific trend where higher rates of depression correlated with reduced recognition of sad faces (18). Thirteen studies also collected measures of anxiety, two of which also explored OCD symptomology through validated measures. All studies reported significantly higher anxiety as well as OCD traits in high eating pathology groups; however, no relationships between anxiety, OCD symptoms and ER performance were identified.

Eight studies measured alexithymia using the Toronto Alexithymia Scale (TAS; Bagby, Parker, & Taylor, 1994). As expected, most studies identified higher rates of alexithymia among ED groups (3, 7, 8, 9, 15, 18). One study found a significant relationship between alexithymia and ER ability (6), while another study identified a specific relationship between alexithymia score and reduced recognition of neutral faces (3). The remaining six studies exploring this construct did not find any relationship (7, 8, 9, 15, 18, 19). Overall, the majority of studies accounted for co-occurring symptoms, and there is little evidence from the present review that difficulties with ER in high eating pathology groups are due to co-morbid factors associated with ED.

Experimental Measures

The measures used in each study varied considerably. Five studies used the mentalising task Reading the Mind in the Eyes (RME, Baron-Cohen et al., 1997). This task does not measure performance on individual emotions and instead asks participants to attribute internal emotional states to pictures of eyes (2, 5, 10, 13, 16). Four of five studies using the RME reported global differences in performance among groups, while the remaining study (16) reported specific differences in recognising negative emotions. The remaining fifteen papers used alterations of previous designs or study-specific bespoke tasks based on the emotion recognition task (ERT) initially described by Ekman and Friesen (1976) (1, 3, 4, 6, 7, 8, 9, 11, 12, 14, 15, 17, 18, 19, 20). Facial stimuli in bespoke tasks were taken from validated sources; however, it is difficult to comment on the reliability of these experimental tasks as they have each been constructed by their respective authors (8, 12, 15). As the majority of experimental ER tasks in the studies reviewed here are bespoke, comparability across studies is restricted. Indeed, with two exceptions (4, 17), each of the tasks differed in their procedure such that it is not possible to comment on replicability and direct comparisons are challenging (Paiva-Silva, Pontes, Aguiar, & de Souza, 2016). Five studies reported skewed data on ER measures due to ceiling effects (8, 9, 10, 12, 16) suggestive of poor task sensitivity.

All but two studies in this review used forced-choice labelling of emotions. The two tasks that did not follow this design required yes or no responses to specific emotion identification questions (6, 11) and failed to identify ER differences across groups. The number of emotions assessed ranged from two to nine with most tests

including six. The number of emotions measured did not influence the likelihood of finding significant differences between groups. However, one study which included only two emotions failed to find any differences between groups (11). Fourteen studies used static images of faces displaying an emotive expression. Six studies used morphed stimuli in their bespoke experimental design that either altered the intensity of the emotional expression or presented two emotion expressions together (4, 5, 12, 17, 18, 20). The likelihood of significant differences between groups did not differ between static and morphed stimulus designs. Indeed, two studies using morphed stimuli only found a significant difference between groups on less ambiguous stimuli (8, 12).

Finally, one study failed to find differences between groups using only male face stimuli (6), while another study identified that performance only differed in the high eating pathology group when observing female face stimuli (2). Tasks such as the RME (2, 5, 10, 13, 16) as well as several ERT studies (11, 20) did not have a time limit for either stimulus presentation or participant response. Excluding the RME tasks, tasks that did not have a time limit on stimuli presentation did not find differences in ER between groups, suggesting this could be a significant factor in determining performance on such tasks. These tasks along with other designs with extended display latencies of 10 seconds or more (1, 7, 14) cannot be considered ecologically valid. Most other study designs (6, 4, 9, 15, 17, 18, 19) displayed images for two seconds or less which is more representative of real-life social interaction. Finally, only seven studies in this review calculated misclassification data (e.g. the number of times anger was selected for the wrong face stimuli; 3, 4, 8, 12, 14, 17, 18). All the papers that included misclassification data reported significant ER difference between groups suggesting

there may be a role of attentional biases in ER and eating pathology. In summary, ER task design has a significant impact on the performance of participants in high and low eating pathology groups. Parameters that promote greater task sensitivity such as forced labelling response, shorter presentation time of stimuli, misclassification data and stimulus factors may all contribute towards more accurate detection of ER performance in eating disordered populations.

Discussion

This paper aimed to critically review whether women with high eating pathology exhibit deficits when recognising emotions in faces. Across twenty studies, global deficits were present in half of the studies and specific differences were present in a further twelve. Only three studies failed to demonstrate any difference in ER between groups; however, these tasks had particularly weak methodologies. There did not appear to be a difference between the magnitude of difference in ER deficit and the eating pathology measured (e.g. AN, BN). Collectively, the data suggest there are fundamental differences in ER in populations with high levels of eating pathology; however, the exact nature of this difficulty remains unclear.

Difficulties recognising disgust were present in several studies of AN which is unexpected as AN is often associated with greater disgust sensitivity, particularly concerning food and body image (Aharoni & Hertz, 2011; Chu, Bodell, Ribiero, & Joiner, 2015). However, contemporary models of AN suggest that poor introspective ability observed in high eating pathology groups may relate to difficulty in interpreting complex negative emotions such as disgust both internally and in others (Moncrieff-Boyd, Bryne,

& Nunn, 2013). Similarly, reduced disgust recognition is integral to neurobiological hypotheses such as the dysfunctional insula model that associates abnormal insula activity in AN with reduced processing of disgust (Nunn, Frampton, Gordon, & Lask, 2008).

Another notable trend was a specific cognitive bias towards faces depicting anger. This finding is consistent with research that suggests that individuals with ED have different attentional responses to angry faces (Harrison et al., 2010; Pringle, Harmer, & Cooper, 2010). Indeed, hypervigilance to threatening stimuli is often observed in eating disorders (Gilon et al., 2018), potentially making angry faces easier to detect in ER paradigms (Harrison et al., 2010). Furthermore, Fox et al (2013) suggest that feelings of anger and disgust may have become 'coupled' in individuals with ED, making them difficult to distinguish. Other theories propose that internal avoidance of emotional experience is responsible for poor recognition of emotion in others, as similar brain regions are recruited in both the experience and interpretation of emotions (Calder & Young, 2005).

The role of clinical and co-morbid factors was also examined to better understand the variability across studies. The review did not find evidence that ER performance was related to a state of starvation (Harrison et al., 2010) and diminished ER was observed in at-risk groups as well as clinical populations thus suggesting ER may be a trait of eating pathology that exists independently of illness. Similarly, contrary to previous reports (Montagne et al., 2006) the present review suggests that highly comorbid factors including depression (O'Brien & Vincent, 2003) and anxiety (Goddard & Treasure, 2013; Swinbourne et al., 2012) are not associated with ER in high eating

pathology groups. Finally, alexithymia has been argued to contribute towards ER deficits in EDs (Bird & Cook, 2013); however, most of the papers reviewed here do not support this conclusion. Overall, the reviewed studies support the position that observed differences in ER in high eating pathology populations are not associated with comorbidity or state or severity of illness.

One factor that does appear to contribute towards the variability in findings is the limitations of the experimental tasks used to measure ER (Paiva-Silva et al., 2016). For example, the Reading the Mind in the Eyes task (RME, Baron-Cohen et al., 2001) used in five reviewed studies may have insufficient ecological validity as it only displays eyes and has no presentation time-limit. However, , studies using this task did report a significant difference between high and low eating pathology groups, and these are the only studies that can be reliably compared due to their identical task procedure.

Most other tasks reviewed used previously validated static face stimuli. However, the ecological validity of these images is limited as they are not representative of the complexity of facial emotion displays in real life (Chafi, 2012). Indeed, individuals who are motivated may be able to succeed at this task without having an internal working model of the meaning of the emotions they recognise visually in faces (Kessler et al., 2006). Several studies did overcome this difficulty by presenting more sensitive stimuli that had been morphed (Rosenberg, McDonald, Dethier, Kessels, & Westbrook, 2014). Interestingly, while such studies were the minority in this review, they appeared to show the same level of deficit in ER performance in high eating pathology groups as static stimuli designs. Indeed, two designs using morphed stimuli only found differences

between groups when stimuli were less ambiguous suggesting that morphed face stimuli are equally challenging for clinical and control individuals.

While it may be advantageous that literature in this field has not relied on a single approach to measure ER, bespoke and study-specific experimental designs make comparisons between studies challenging. Equally, many different task parameters may have contributed towards the inconsistency of findings. For example, extended display latencies are less representative of real-life facial expressions which are often brief (Surguladze et al., 2004) and studies with longer latencies were less likely to find ER differences among groups. Similarly, while most tasks used forced-choice recognition, two studies used a yes/no response to emotional stimuli and failed to identify differences between groups. Finally, while face stimuli were all taken from validated sources, face characteristics including gender and age can impact ER, and unreported image dimensions may contribute towards the diversity of results (Parmley & Cunningham, 2014). In sum, while a full consensus was not found, most of the studies in this field support the position that individuals with high eating pathology have global or specific differences in ER ability compared to their low eating pathology counterparts. Effect sizes are predominantly medium to large suggesting that differences in ER are clinically significant; however, it is important to note that a quarter of the studies reviewed were underpowered. Finally, the review also concludes that variability observed between studies are less likely to be due to the clinical population or co-morbidities of the groups measured, but rather the diversity in the methodological design of ER tasks.

Limitations and Future Directions

The present review has several limiting factors that need to be acknowledged. First, all papers used a female sample such that the present review can only be generalised to adult females. Neurological evidence suggests that the way men and women process emotions is considerably different (Whittle, Yucel, Yap, & Allen, 2011) and some studies suggest that women are better at detecting subtle emotion expressions than men (Hoffman, Kessler, Eppel, Rukavina, & Traue, 2010). It is possible, therefore, that the deficits observed in these samples may be more significant in male samples (Hall & Matsumoto, 2004). Future reviews may wish to consider the impact of gender on the conclusions drawn here.

Secondly, the present review only considered emotion recognition in still images of facial stimuli. Dynamic stimuli presented through video clips present an entirely different means of measuring ER that can overcome some of the limitations concerning ecological validity evident in the measures reported in this review (Gutiérrez-Maldonado, Rus-Calafell, & González-Conde, 2014). Indeed, there is evidence to suggest that motion cues significantly improve recognition of subtle emotions (Bould & Morris, 2008) and different neural networks are employed when recognising static and dynamic facial expressions (Kilts, Egan, Gideon, Ely, & Hoffman, 2003).

Alternative means of measuring ER were also not included in the present review. One emerging field of ER research is neuroimaging methods including electroencephalography (EEG; Kirihara et al., 2012), eye-tracking (Calvo & Nummenmaa, 2009) and functional magnetic resonance imaging (fMRI; Johnston, Stokanov, Devir, & Schall., 2005). It is acknowledged that different emotional

expressions recruit distinct neural substrates. For example, disgust processing has been found to be related to the insula (Aleman & Swart, 2008), while anger expressions are processed by the cingulate cortex (Blair, Morris, Frith, Perrett, & Dolan, 1999). Understanding the brain-behaviour relationship in this domain can help inform and test new models of eating pathology such as the insula hypothesis of AN (Nunn et al., 2008). However, such measures are challenging to access, and the neurobiological correlates of facial processing can be difficult to compare due to differences in instruments used.

All of the studies reviewed here were subjected to an assessment of their quality through the Quality Assessment Tool of Quantitative Studies (QATQS). This tool is designed to measure quantitative studies and rewards gold-standard research designs such as randomised control-trials. As such, the tool is less sensitive to other research methodologies and has limited the variability in quality ratings across the reviewed studies. For example, as all designs included were case controls, they did not differ on their scores regarding study design, blinding of participants or withdrawal rates. While this limits the function of this tool to some degree it does not limit the conclusions of this review as the methodological design of identified papers were examined in detail. Future research may employ a broader range of quality assessment methods to overcome the limitation of tools such as the QATQS.

Conclusion

The present review explored whether there are meaningful differences in ER ability between women with high eating pathology and controls. The review identified twenty papers, of which half reported global ER deficits in high eating pathology groups.

Only three studies found no effect and there were considerable limitations in the experimental design of these papers. Thus, differences in task design may account for the variation in trends observed in the recognition of specific emotions such as anger and disgust. Several confounding variables including co-morbidity, illness type and illness severity do not convincingly account for the deficits in ER observed between groups. While most papers reviewed support the position that ER differences exist in eating pathology, heterogeneity in study design prevent us from drawing a satisfying conclusion as to the exact nature of this difficulty. Encouragingly, there are many new directions that this field of research may take to further our understanding of eating pathology and ER.

References

Adenzato, M., Todisco, P., & Ardito, R. B. (2012). Social cognition in anorexia nervosa: evidence of preserved theory of mind and impaired emotional functioning. *PloS One*, *7*, e44414.

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, *3*, 469-479.

Aharoni, R., & Hertz, M. M. (2012). Disgust sensitivity and anorexia nervosa. *European Eating Disorders Review*, *20*, 106-110.

Aleman, A., & Swart, M. (2008). Sex differences in neural activation to facial expressions denoting contempt and disgust. *PloS One*, *3*, e3622.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub.

Armijo-Olivo, S., Stiles, C. R., Hagen, N. A., Biondo, P. D., & Cummings, G. G. (2012). Assessment of study quality for systematic reviews: a comparison of the

Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *Journal of Evaluation in Clinical Practice*, *18*, 12-18.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268.

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*, 23-32.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child psychology and Psychiatry*, *38*, 813-822.

Bastiaansen, J. A., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 2391-2404.

Bekker, M. H., & Spoor, S. T. (2008). Emotional inhibition, health, gender, and eating disorders: The role of (over) sensitivity to others. *Emotion Regulation*, *24,* 170-183.

Berg, K. C., Peterson, C. B., Frazier, P., & Crow, S. J. (2012). Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. *International Journal of Eating Disorders*, *45*, 428-438.

Bird, G., & Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry*, *3*, e285.

Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain*, *122*, 883-893.

Blinder, B. J., Cumella, E. J., & Sanathara, V. A. (2006). Psychiatric comorbidities of female inpatients with eating disorders. *Psychosomatic Medicine*, *68*, 454-462.

Bould, E., & Morris, N. (2008). Role of motion signals in recognizing subtle facial expressions of emotion. *British Journal of Psychology*, *99*, 167-189.

Bourke, C., Douglas, K., & Porter, R. (2010). Processing of facial emotion expression in major depression: A review. *Australian and New Zealand Journal of Psychiatry, 44,* 681– 696.

Brewer, R., Cook, R., Cardi, V., Treasure, J., & Bird, G. (2015). Emotion recognition deficits in eating disorders are explained by co-occurring alexithymia. *Royal Society Open Science*, *2*, 140382.

Brockmeyer, T., Skunde, M., Wu, M., Bresslein, E., Rudofsky, G., Herzog, W., & Friederich, H. C. (2014). Difficulties in emotion regulation across the spectrum of eating disorders. *Comprehensive Psychiatry*, *55*, 565-571.

Caglar-Nazali, H. P., Corfield, F., Cardi, V., Ambwani, S., Leppanen, J., Olabintan, O., ... Treasure, J. (2014). A systematic review and meta-analysis of 'Systems for Social Processes' in eating disorders. *Neuroscience & Biobehavioral Reviews, 42,* 55–92.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*, 641-651.

Calvo, M. G., & Nummenmaa, L. (2009). Eye-movement assessment of the time course in facial expression recognition: Neurophysiological implications. *Cognitive, Affective, & Behavioral Neuroscience*, *9*, 398-411.

Cardi, V., Corfield, F., Leppanen, J., Rhind, C., Deriziotis, S., Hadjimichalis, A., ... & Treasure, J. (2015). Emotional processing, recognition, empathy and evoked facial expression in eating disorders: an experimental study to map deficits in social cognition. *PLoS One*, *10*, e0133827.

Cardi, V., Mallorquí-Bagué, N., Albano, G., Monteleone, A. M., Fernandez-Aranda, F., & Treasure, J. (2018). Social difficulties as risk and maintaining factors in anorexia nervosa: a mixed-method investigation. *Frontiers in Psychiatry*, *9*, 12.

Castro, L., Davies, H., Hale, L., Surguladze, S., & Tchanturia, K. (2010). Facial affect recognition in anorexia nervosa: is obsessionality a missing piece of the puzzle?. *Australian & New Zealand Journal of Psychiatry*, *44*, 1118-1125.

Chafi, A. (2012). Three patterns of motion which change the perception of emotional faces. *Psychology, 3*, 82– 89.

Chu, C., Bodell, L. P., Ribeiro, J. D., & Joiner, T. E. (2015). Eating disorder symptoms and suicidal ideation: the moderating role of disgust. *European Eating Disorders Review*, *23*, 545-552.

Cooper, M. J., Wells, A., & Todd, G. (2004). A cognitive model of bulimia nervosa. *British Journal of Clinical Psychology*, *43*, 1-16.

Dalili, M. N., Penton-Voak, I. S., Harmer, C. J., & Munafò, M. R. (2015). Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychological Medicine*, *45*, 1135-1144.

Dapelo, M. M., Surguladze, S., Morris, R., & Tchanturia, K. (2016). Emotion recognition in blended facial expressions in women with anorexia nervosa. *European Eating Disorders Review*, *24*, 34-42.

Dapelo, M. M., Surguladze, S., Morris, R., & Tchanturia, K. (2017). Emotion Recognition in Face and Body Motion in Bulimia Nervosa. *European Eating Disorders Review*, *25*, 595-600.

Dellava, J. E., Thornton, L. M., Hamer, R. M., Strober, M., Plotnicov, K., Klump, K. L., … Bulik, C. M. (2010). Childhood anxiety associated with low BMI in women with anorexia nervosa. *Behaviour Research and Therapy, 48*, 60–67.

Demenescu, L. R., Kortekaas, R., den Boer, J. A., & Aleman, A. (2010). Impaired attribution of emotion to facial expressions in anxiety and major depression. *PloS One*, *5*, e15058.

Ekman P., Friesen W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personal Social Psychology, 17,* 124–129.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*, 363-370.

Montagne, B., Schutters, S., Westenberg, H. G., van Honk, J., Kessels, R. P., & de Haan, E. H. (2006). Reduced sensitivity in the recognition of anger and disgust in social anxiety disorder. *Cognitive Neuropsychiatry*, *11*, 389-401.

Fox, J. R., Smithson, E., Baillie, S., Ferreira, N., Mayr, I., & Power, M. J. (2013). Emotion coupling and regulation in anorexia nervosa. *Clinical Psychology & Psychotherapy*, *20*, 319-333.

Frith, U., & Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *358*, 459–473.

Gilboa-Schechtman, E., Avnon, L., Zubery, E., & Jeczmien, P. (2006). Emotional processing in eating disorders: specific impairment or general distress related deficiency?. *Depression and Anxiety*, *23*, 331-339.

Gilon, M. T., Hamdan, S., Bar-Haim, Y., Lazarov, A., Enoch-Levy, A., Dubnov-Raz, G., ... & Stein, D. (2018). Different attention bias patterns in anorexia nervosa restricting and binge/purge types. *European Eating Disorders Review,* 1-9.

Goddard, E., & Treasure, J. (2013). Anxiety and social-emotional processing in eating disorders: examination of family trios. *Cognitive Therapy and Research, 37*, 890-904.

Gramaglia, C., Ressico, F., Gambaro, E., Palazzolo, A., Mazzarino, M., Bert, F., ... & Zeppegno, P. (2016). Alexithymia, empathy, emotion identification and social inference in anorexia nervosa: A case-control study. *Eating Behaviors*, *22*, 46-50.

Gutiérrez-Maldonado, J., Rus-Calafell, M., & González-Conde, J. (2014). Creation of a new set of dynamic virtual reality faces for the assessment and training of facial emotion recognition ability. *Virtual Reality*, *18*, 61-71.

Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, *4*, 201.

Harrison, A., Sullivan, S., Tchanturia, K., & Treasure, J. (2009). Emotion recognition and regulation in anorexia nervosa. *Clinical Psychology & Psychotherapy*, *16*, 348-356.

Harrison, A., Sullivan, S., Tchanturia, K., & Treasure, J. (2010). Emotional functioning in eating disorders: attentional bias, emotion recognition and emotion regulation. *Psychological Medicine*, *40*, 1887-1897.

Harrison, A., Tchanturia, K., & Treasure, J. (2010). Attentional bias, emotion recognition, and emotion regulation in anorexia: state or trait?. *Biological Psychiatry*, *68*, 755-761.

Haynos, A. F., & Fruzzetti, A. E. (2011). Anorexia nervosa as a disorder of emotion dysregulation: Evidence and treatment implications. *Clinical Psychology: Science and Practice*, *18*, 183-202.

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica*, *135,* 278-283.

Holliday, J., Uher, R., Landau, S., Collier, D., & Treasure, J. (2006). Personality pathology among individuals with a lifetime history of anorexia nervosa. *Journal of Personality Disorders*, *20*, 417-430.

Horning, S. M., Cornwell, R. E., & Davis, H. P. (2012). The recognition of facial expressions: an investigation of the influence of age and cognition. *Aging, Neuropsychology, and Cognition*, *19*, 657-676.

Isaacowitz, D. M., Löckenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging*, *22*, 147.

Jänsch, C., Harmer, C., & Cooper, M. J. (2009). Emotional processing in women with anorexia nervosa and in healthy volunteers. *Eating Behaviors*, *10*, 184-191.

Johnston, P. J., Stojanov, W., Devir, H., & Schall, U. (2005). Functional MRI of facial emotion recognition deficits in schizophrenia and their electrophysiological correlates. *European Journal of Neuroscience*, *22*, 1221-1232.

Jones, L., Harmer, C., Cowen, P., & Cooper, M. (2008). Emotional face processing in women with high and low levels of eating disorder related symptoms. *Eating Behaviors*, *9*, 389-397.

Kaye, W. H., Wierenga, C. E., Bailer, U. F., Simmons, A. N., & Bischoff-Grethe, A. (2013). Nothing tastes as good as skinny feels: the neurobiology of anorexia nervosa. *Trends in Neurosciences, 36*, 110–120.

Kenyon, M., Samarawickrema, N., DeJong, H., Van den Eynde, F., Startup, H., Lavender, A., ... & Schmidt, U. (2012). Theory of mind in bulimia nervosa. *International Journal of Eating Disorders*, *45*, 377-384.

Kessler, H., Schwarze, M., Filipic, S., Traue, H. C., & von Wietersheim, J. (2006). Alexithymia and facial emotion recognition in patients with eating disorders. *International Journal of Eating Disorders*, *39*, 245-251.

Kirihara, K., Kasai, K., Tada, M., Nagai, T., Kawakubo, Y., Yamasaki, S., ... & Araki, T. (2012). Neurophysiological impairment in emotional face processing is associated with low extraversion in schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *37*, 270-275.

Kucharska-Pietura, K., Nikolaou, V., Masiak, M., & Treasure, J. (2004). The recognition of emotion in the faces and voice of anorexia nervosa*. International Journal of Eating Disorders, 35*, 42-47.

Kühnpast, N., Gramann, K., & Pollatos, O. (2012). Electrophysiologic evidence for multilevel deficits in emotional face processing in patients with bulimia nervosa. *Psychosomatic Medicine*, *74*, 736-744.

Ladegaard, N., Videbech, P., Lysaker, P. H., & Larsen, E. R. (2016). The course of social cognitive and metacognitive ability in depression: Deficit are only partially normalized after full remission of first episode major depression. *British Journal of Clinical Psychology*, *55*, 269-286.

Lane, R. D., Sechrest, L., Riedel, R., Shapiro, D. E., & Kaszniak, A. W. (2000). Pervasive emotion recognition deficit common to alexithymia and the repressive coping style. *Psychosomatic Medicine, 62*, 492-501.

Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology*, *6*, 761.

Legenbauer, T., Vocks, S., & Rüddel, H. (2008). Emotion recognition, emotional awareness and cognitive bias in individuals with bulimia nervosa. *Journal of Clinical Psychology*, *64*, 687-702.

Lulé, D., Schulze, U. M., Bauer, K., Schöll, F., Müller, S., Fladung, A. K., & Uttner, I. (2014). Anorexia nervosa and its relation to depression, anxiety, alexithymia and emotional processing deficits. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *19*, 209-216.

Marcon, D. T. (2017). *Emotional facial expression recognition and emotion regulation in undergraduate students with disordered eating* (Doctoral dissertation, Laurentian University of Sudbury).

Marsden, P., Karagianni, E., & Morgan, J. F. (2007). Spirituality and clinical care in eating disorders: a qualitative study. *The International Journal of Eating Disorders, 40,* 7–12.

McAuley, L., Tugwell, P., & Moher, D. (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses?. *The Lancet*, *356*, 1228-1231.

McCormick, L. M., Keel, P. K., Brumm, M. C., Bowers, W., Swayze, V., Andersen, A., & Andreasen, N. (2008). Implications of starvation-induced change in right dorsal anterior cingulate volume in anorexia nervosa*. International Journal of Eating Disorders, 41,* 602-610.

Medina-Pradas, C., Blas Navarro, J., Álvarez-Moya, E. M., Grau, A., & Obiols, J. E. (2012). Emotional theory of mind in eating disorders. *International Journal of Clinical and Health Psychology*, *12*, 189-202.

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., Cheraghi-Sohi, S., Stephens, K., … Reindeers, A. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research*, *14*, 579.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*, 1.

Moncrieff-Boyd, J., Byrne, S., & Nunn, K. (2014). Disgust and anorexia nervosa: confusion between self and non-self. *Advances in Eating Disorders: Theory, Research and Practice*, *2*, 4-18.

Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test (NART)*. Windsor: Nfer-Nelson.

Nowakowski, M. E., McFarlane, T., & Cassin, S. (2013). Alexithymia and eating disorders: a critical review of the literature. *Journal of Eating Disorders*, *1*, 21.

Nunn, K., Frampton, I., Gordon, I., & Lask, B. (2008). The fault is not in her parents but in her insula—a neurobiological hypothesis of anorexia nervosa. *European Eating Disorders Review*, *16*, 355-360.

O'Brien, K. M., & Vincent, N. K. (2003). Psychiatric comorbidity in anorexia and bulimia nervosa: nature, prevalence, and causal relationships. *Clinical Psychology Review, 23,* 57-74.

Oldershaw, A., Hambrook, D., Stahl, D., Tchanturia, K., Treasure, J., & Schmidt, U. (2011). The socio-emotional processing stream in anorexia nervosa. *Neuroscience & Biobehavioral Reviews*, *35*, 970-988.

Oldershaw, A., Hambrook, D., Tchanturia, K., Treasure, J., & Schmidt, U. (2010). Emotional theory of mind and emotional awareness in recovered anorexia nervosa patients. *Psychosomatic Medicine*, *72*, 73-79.

Ochsner, K. N. (2008). The social-emotional processing stream: five core constructs and their translational potential for schizophrenia and beyond. *Biological Psychiatry*, *64*, 48-61.

Paiva-Silva, A. I. D., Pontes, M. K., Aguiar, J. S. R., & de Souza, W. C. (2016). How do we evaluate facial emotion recognition?. *Psychology & Neuroscience*, *9*, 153.

Parmley, M., & Cunningham, J. G. (2014). She looks sad, but he looks mad: The effects of age, gender, and ambiguity on emotion perception. *The Journal of Social Psychology, 154,* 323-338.

Pollatos, O., Herbert, B. M., Schandry, R., & Gramann, K. (2008). Impaired central processing of emotional faces in anorexia nervosa. *Psychosomatic Medicine*, *70*, 701-708.

Pringle, A., Harmer, C. J., & Cooper, M. J. (2010). Investigating vulnerability to eating disorders: biases in emotional processing. *Psychological Medicine*, *40*, 645-655.

Ridout, N., Wallis, D. J., Autwal, Y., & Sellis, J. (2012). The influence of emotional intensity on facial emotion recognition in disordered eating. *Appetite*, *59*, 181-186.

Rosenberg, H., McDonald, S., Dethier, M., Kessels, R. P., & Westbrook, R. F. (2014). Facial emotion recognition deficits following moderate–severe traumatic brain injury (TBI): Re-examining the valence effect and the role of emotion intensity. *Journal of the International Neuropsychological Society*, *20*, 994-1003.

Rothschild-Yakar, L., Eviatar, Z., Shamia, A., & Gur, E. (2011). Social cognition in eating disorders: encoding and representational processes in binging and purging patients. *European Eating Disorders Review*, *19*, 75-84.

Russell, T. A., Schmidt, U., Doherty, L., Young, V., & Tchanturia, K. (2009). Aspects of social cognition in anorexia nervosa: affective and cognitive theory of mind. *Psychiatry Research*, *168*, 181-185.

Schmidt, U., & Treasure, J. (2006). Anorexia nervosa: Valued and visible. A cognitive-interpersonal maintenance model and its implications for research and practice. *British Journal of Clinical Psychology*, *45*, 343-366.

Seiferth, N. Y., Pauly, K., Kellermann, T., Shah, N. J., Ott, G., Herpertz-Dahlmann, B.,... Habel, U. (2009). Neuronal correlates of facial emotion discrimination in early onset schizophrenia. *Neuropsychopharmacology, 34,* 477– 487.

Sharpe, E., Wallis, D. J., & Ridout, N. (2016). The influence of variations in eating disorder-related symptoms on processing of emotional faces in a non-clinical female sample: An eye-tracking study. *Psychiatry Research*, *240*, 321-327.

Speranza, M., Loas, G., Wallier, J., & Corcos, M. (2007). Predictive value of alexithymia in patients with eating disorders: A 3-year prospective study. *Journal of Psychosomatic Research, 63*, 365–371.

Surguladze, S. A., Young, A. W., Senior, C., Brébion, G., Travis, M. J., & Phillips, M. L. (2004). Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. *Neuropsychology*, *18*, 212.

Swinbourne, J., Hunt, C., Abbott, M., Russell, J., St Clare, T., & Touyz, S. (2012). The comorbidity between eating disorders and anxiety disorders: Prevalence in an eating disorder sample and anxiety disorder sample. *Australian & New Zealand Journal of Psychiatry*, *46*, 118-131.

Taylor, G. J. (1994). The alexithymia construct: Conceptualization, validation, and relationship with basic dimensions of personality. *New Trends in Experimental & Clinical Psychiatry, 10*, 61-74.

Tchanturia, K., Davies, H., Harrison, A., Fox, J. R., Treasure, J., & Schmidt, U. (2012). Altered social hedonic processing in eating disorders. *International Journal of Eating Disorders*, *45*, 962-969.

Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews on Evidence‐Based Nursing*, *1*, 176-184.

Treasure, J., & Schmidt, U. (2013). The cognitive-interpersonal maintenance model of anorexia nervosa revisited: a summary of the evidence for cognitive, socio-emotional and interpersonal predisposing and perpetuating factors. *Journal of Eating Disorders*, *1*, 13.

Weightman, M. J., Air, T. M., & Baune, B. T. (2014). A review of the role of social cognition in major depressive disorder. *Frontiers in Psychiatry*, *5*, 179.

Whittle, S., Yücel, M., Yap, M. B., & Allen, N. B. (2011). Sex differences in the neural correlates of emotion: evidence from neuroimaging. *Biological Psychology*, *87*, 319-333.

Witztum, E., Latzer, Y., & Stein, D. (2008). Anorexia nervosa and bulimia nervosa as idioms of distress: From the historical background to current formulations. *International Journal of Child Adolescent Health*, *1*, 283-294.

Appendices

Appendix A – Quality Assessment tool for Quantitative Studies

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?

1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?

1 80 - 100% agreement
2 60 – 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.

No Yes

If Yes, was the method of randomization described? (See dictionary)

No Yes

If Yes, was the method appropriate? (See dictionary)

No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention? 1 Yes

2 No
3 Can't tell

The following are examples of confounders:

1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1 80 – 100% (most)
2 60 – 79% (some)
3 Less than 60% (few or none)
4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?

1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question? 1 Yes

2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?

1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)

community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)

community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?

1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

1 Yes
2 No
3 Can't tell

GLOBAL RATING

GLOBAL RATING FOR THIS PAPER (circle one):

1	STRONG	(no WEAK ratings)
2	MODERATE	(one WEAK rating)
3	WEAK	(two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

No Yes

If yes, indicate the reason for the discrepancy

1 Oversight
2 Differences in interpretation of criteria
3 Differences in interpretation of study

Final decision of both reviewers (circle one):

1. **STRONG**
2. **MODERATE**
3. **WEAK**

Appendix B – Quality Assessment tool for Quantitative Studies Dictionary

Quality Assessment Tool for Quantitative Studies Dictionary

The purpose of this dictionary is to describe items in the tool thereby assisting raters to score study quality. Due to under-reporting or lack of clarity in the primary study, raters will need to make judgements about the extent that bias may be present. When making judgements about each component, raters should form their opinion based upon information contained in the study rather than making inferences about what the authors intended. Mixed methods studies can be quality assessed using this tool with the quantitative component of the study.

A) SELECTION BIAS

(Q1) Participants are more likely to be representative of the target population if they are randomly selected from a comprehensive list of individuals in the target population (score very likely). They may not be representative if they are referred from a source (e.g. clinic) in a systematic manner (score somewhat likely) or self-referred (score not likely).

(Q2) Refers to the % of subjects in the control and intervention groups that agreed to participate in the study before they were assigned to intervention or control groups.

B) STUDY DESIGN

In this section, raters assess the likelihood of bias due to the allocation process in an experimental study. For observational studies, raters assess the extent that assessments of exposure and outcome are likely to be independent. Generally, the type of design is a good indicator of the extent of bias. In stronger designs, an equivalent control group is present and the allocation process is such that the investigators are unable to predict the sequence.

Randomized Controlled Trial (RCT)

An experimental design where investigators randomly allocate eligible people to an intervention or control group. A rater should describe a study as an RCT if the randomization sequence allows each study participant to have the same chance of receiving each intervention and the investigators could not predict which intervention

was next. If the investigators do not describe the allocation process and only use the words ‘random’ or ‘randomly’, the study is described as a controlled clinical trial.

See below for more details.

Was the study described as randomized?

Score YES, if the authors used words such as random allocation, randomly assigned, and random assignment. Score NO, if no mention of randomization is made.

Was the method of randomization described?

Score YES, if the authors describe any method used to generate a random allocation sequence.

Score NO, if the authors do not describe the allocation method or describe methods of allocation such as alternation, case record numbers, dates of birth, day of the week, and any allocation procedure that is entirely transparent before assignment, such as an open list of random numbers of assignments.

If NO is scored, then the study is a controlled clinical trial.
Was the method appropriate?

Score YES, if the randomization sequence allowed each study participant to have the same chance of receiving each intervention and the investigators could not predict which intervention was next. Examples of appropriate approaches include assignment of subjects by a central office unaware of subject characteristics, or sequentially numbered, sealed, opaque envelopes.

Score NO, if the randomization sequence is open to the individuals responsible for recruiting and allocating participants or providing the intervention, since those individuals can influence the allocation process, either knowingly or unknowingly.

If NO is scored, then the study is a controlled clinical trial.

Controlled Clinical Trial (CCT)
An experimental study design where the method of allocating study subjects to intervention or control groups is open to individuals responsible for recruiting subjects or providing the intervention. The method of allocation is transparent before assignment, e.g. an open list of random numbers or allocation by date of birth, etc.

Cohort analytic (two group pre and post)
An observational study design where groups are assembled according to whether or not exposure to the intervention has occurred. Exposure to the intervention is not under the control of the investigators. Study groups might be non-equivalent or not comparable on some feature that affects outcome.

Case control study
A retrospective study design where the investigators gather ‘cases’ of people who already have the outcome of interest and ‘controls’ who do not. Both groups are then questioned or their records examined about whether they received the intervention exposure of interest.

Cohort (one group pre + post (before and after)

The same group is pretested, given an intervention, and tested immediately after the intervention. The intervention group, by means of the pretest, act as their own control group.

Interrupted time series
A study that uses observations at multiple time points before and after an intervention (the 'interruption'). The design attempts to detect whether the intervention has had an effect significantly greater than any underlying trend over time. Exclusion: Studies that do not have a clearly defined point in time when the intervention occurred and at least three data points before and three after the intervention

Other:
One time surveys or interviews

C) CONFOUNDERS

By definition, a confounder is a variable that is associated with the intervention or exposure and causally related to the outcome of interest. Even in a robust study design, groups may not be balanced with respect to important variables prior to the intervention. The authors should indicate if confounders were controlled in the design (by stratification or matching) or in the analysis. If the allocation to intervention and control groups is randomized, the authors must report that the groups were balanced at baseline with respect to confounders (either in the text or a table).

BLINDING

(Q1) Assessors should be described as blinded to which participants were in the control and intervention groups. The purpose of blinding the outcome assessors (who might also be the care providers) is to protect against detection bias.

(Q2) Study participants should not be aware of (i.e. blinded to) the research question. The purpose of blinding the participants is to protect against reporting bias.

D) DATA COLLECTION METHODS

Tools for primary outcome measures must be described as reliable and valid. If 'face' validity or 'content' validity has been demonstrated, this is acceptable. Some sources from which data may be collected are described below:

Self reported data includes data that is collected from participants in the study (e.g. completing a questionnaire, survey, answering questions during an interview, etc.).

Assessment/Screening includes objective data that is retrieved by the researchers. (e.g. observations by investigators).

Medical Records/Vital Statistics refers to the types of formal records used for the extraction of the data.

Reliability and validity can be reported in the study or in a separate study. For example, some standard assessment tools have known reliability and validity.

E) WITHDRAWALS AND DROP-OUTS

Score **YES** if the authors describe BOTH the numbers and reasons for withdrawals and drop-outs. Score **NO** if either the numbers or reasons for withdrawals and drop-outs are not reported.

Score **NOT APPLICABLE** if the study was a one-time interview or survey where there was not follow-up data reported.

The percentage of participants completing the study refers to the % of subjects remaining in the study at the final data collection period in all groups (i.e. control and intervention groups).

F) INTERVENTION INTEGRITY

The number of participants receiving the intended intervention should be noted (consider both frequency and intensity). For example, the authors may have reported that at least 80 percent of the participants received the complete intervention. The authors should describe a method of measuring if the intervention was provided to all participants the same way. As well, the authors should indicate if subjects received an unintended intervention that may have influenced the outcomes. For example, co-intervention occurs when the study group receives an additional intervention (other than that intended). In this case, it is possible that the effect of the intervention may be over-estimated. Contamination refers to situations where the control group accidentally receives the study intervention. This could result in an under-estimation of the impact of the intervention.

G) ANALYSIS APPROPRIATE TO QUESTION

Was the quantitative analysis appropriate to the research question being asked?

An intention-to-treat analysis is one in which all the participants in a trial are analyzed according to the intervention to which they were allocated, whether they received it or not. Intention-to-treat analyses are favoured in assessments of effectiveness as they mirror the noncompliance and treatment changes that are likely to occur when the intervention is used in practice, and because of the risk of attrition bias when participants are excluded from the analysis.

Component Ratings of Study:

For each of the six components A – F, use the following descriptions as a roadmap.

SELECTION BIAS

Good: The selected individuals are very likely to be representative of the target population (Q1 is 1) **and** there is greater than 80% participation (Q2 is 1).

Fair: The selected individuals are at least somewhat likely to be representative of the target population (Q1 is 1 or 2); **and** there is 60 - 79% participation (Q2 is 2). 'Moderate' may also be assigned if Q1 is 1 or 2 and Q2 is 5 (can't tell).

Poor: The selected individuals are not likely to be representative of the target population (Q1 is 3); **or** there is less than 60% participation (Q2 is 3) **or** selection is not described (Q1 is 4); and the level of participation is not described (Q2 is 5).

DESIGN

Good: will be assigned to those articles that described RCTs and CCTs.

Fair: will be assigned to those that described a cohort analytic study, a case control study, a cohort design, or an interrupted time series.

Weak: will be assigned to those that used any other method or did not state the method used.

CONFOUNDERS

Good: will be assigned to those articles that controlled for at least 80% of relevant confounders (Q1 is 2); **or** (Q2 is 1).

Fair: will be given to those studies that controlled for 60 – 79% of relevant confounders (Q1 is 1) **and** (Q2 is 2).

Poor: will be assigned when less than 60% of relevant confounders were controlled (Q1 is 1) **and** (Q2 is 3) **or** control of confounders was not described (Q1 is 3) **and** (Q2 is 4).

BLINDING

Good: The outcome assessor is not aware of the intervention status of participants (Q1 is 2); **and t**he study participants are not aware of the research question (Q2 is 2).

Fair: The outcome assessor is not aware of the intervention status of participants (Q1 is 2); **or** the study participants are not aware of the research question (Q2 is 2).

Poor: The outcome assessor is aware of the intervention status of participants (Q1 is 1); **and** the study participants are aware of the research question (Q2 is 1); **or b**linding is not described (Q1 is 3 and Q2 is 3).

DATA COLLECTION METHODS

Good: The data collection tools have been shown to be valid (Q1 is 1); **and** the data collection tools have been shown to be reliable (Q2 is 1).

Fair: The data collection tools have been shown to be valid (Q1 is 1); **and** the data collection tools have not been shown to be reliable (Q2 is 2) **or** reliability is not described (Q2 is 3).

Poor: The data collection tools have not been shown to be valid (Q1 is 2) **or** both reliability and validity described (Q1 is 3 and Q2 is 3).

WITHDRAWALS AND DROP-OUTS - a rating of:

Good: will be assigned when the follow-up rate is 80% or greater (Q1 is 1 and Q2 is 1).

Fair: will be assigned when the follow-up rate is 60 – 79% (Q2 is 2) **OR** Q1 is 4 or Q2 is 5.

Poor: will be assigned when a follow-up rate is less than 60% (Q2 is 3) or if the withdrawals and drop-outs were not described (Q1 is No or Q2 is 4).

Not Applicable: if Q1 is 4 or Q2 is 5.

Appendix C - Preparation and Submission Requirements for the European Eating Disorders Review

1. SUBMISSION

Authors should kindly note that submission implies that the content has not been published or submitted for publication elsewhere except as a brief abstract in the proceedings of a scientific meeting or symposium.

Once the submission materials have been prepared in accordance with the Author Guidelines, manuscripts should be submitted online at http://mc.manuscriptcentral.com/erv

2. AIMS AND SCOPE

European Eating Disorders Review provides an international forum for disseminating cutting-edge theoretical and empirical research that significantly advances understanding of the relationship between Eating Disorders and Abnormal Eating/Weight conditions and well-being in humans.

European Eating Disorders Review publishes authoritative and accessible articles, from all over the world, which review or report original research that has implications for the treatment and care of people with eating disorders and obesity, and articles which report innovations and experience in the clinical management of eating disorders. The journal focuses on implications for best practice in diagnosis and treatment. The journal also provides a forum for discussion of the causes and prevention of eating disorders, and related health policy.

Authors may submit original theoretical systematic reviews, methodological, or empirical research articles (7000 words or less) or short communications (3000 words or less). The journal also publishes invited conceptual reviews from leading worldwide researchers in the field of Eating Disorders and/or Obesity. The aims of the journal are to offer a channel of communication between researchers, practitioners, administrators and policymakers who need to report and understand developments in the field of eating disorders.

The journal

- Reports on useful research and experience related to the treatment and prevention of eating disorders in primary care and hospital settings, with special attention to therapy oriented translational research, high quality reviews, clinical trials and pilot innovative therapy approaches.
- Provides information about 'good practice' and systematic reviews.
- Offers a forum for new thinking about the nature, incidence, diagnosis and clinical management of eating disorders (namely anorexia nervosa, bulimia nervosa, binge eating disorders, OSFED and other abnormal eating or feeding behaviors associated with childhood and obesity.

3. MANUSCRIPT CATEGORIES AND REQUIREMENTS

Research articles reporting new research of relevance as set out in the aims and scope should not normally exceed 6000 words (excluding abstract, references, tables or figures), with no more than five tables or illustrations. They should conform to the conventional layout: title page, Abstract, Introduction and Aims, Method, Results, Discussion, Acknowledgements and References. Each of these elements should start on a new page.
Word Limit: 6,000 (excluding abstract, references, tables or figures).

Abstract: 200 words.
References: up to 60.

Review articles: Systematic and meta-analytic review papers are welcomed if they critically review the available literature in a topic than will enhance clinical practice. Articles should have clear focus and enough number of studies should be available for a substantive review paper. Studies that only describe or list previous studies without a critical overview of the literature will not be considered.
Word Limit: 5,000 (excluding abstract, references, tables or figures).
Abstract: 200 words.
References: up to 100.
Figures/Tables: 5 maximum, but should be appropriate to the material covered. Additional tables might be included as supplementary information, if needed. Review articles must follow the PRISMAGuidelines. Authors may want to have a look at the review check lists that reviewers when assessing review articles.

Brief reports should concisely present the essential findings of the author's work and be compromised of the following sections: Abstract, Introduction and Aims, Method, Results, Discussion, and References. Tables and/or figures should be kept to a minimum, in number and size, and only deal with key findings. In some cases authors may be asked to prepare a version of the manuscript with extra material to be included in the online version of the review (as supplementary files). Submissions in this category should not normally exceed 2500 words in length.

Brief reports bring with them a whole host of benefits including: quick and easy submission, administration centralised and reduced and significant decrease in peer review times, first publication priority (this type of manuscript will be published in the next available issue of the journal).

Case Reports The journal does not accept case reports for publication. Authors of case reports are encouraged to submit to the Wiley Open Access journal, Clinical Case Reports www.clinicalcasesjournal.com which aims to directly improve health outcomes by identifying and disseminating examples of best clinical practice.

4. PREPARING THE SUBMISSION

Cover Letters

Cover letters are not mandatory; however, they may be supplied at the author's discretion.

Parts of the Manuscript

The manuscript should be submitted in separate files: title page; main text file; figures.

Main Text File

The text file should be presented in the following order:

i. A short title that contains the major key words. The title should not contain abbreviations (see Wiley's best practice SEO tips);
ii. A short running title of less than 40 characters;
iii. The full names of the authors;
iv. The author's institutional affiliations where the work was conducted, with a footnote for the author's present address if different from where the work was conducted;
v. The corresponding author's contact email address and telephone number;
vi. Acknowledgments;
vii. Conflict of Interest statement (for all authors)
viii. Names and grant numbers of any sources of funding or support in the form of grants, equipment, drugs etc.

Authorship

Please refer to the journal's authorship policy the Editorial Policies and Ethical Considerations section for details on eligibility for author listing eligibility.

Acknowledgments

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section. Financial and material support should also be mentioned. Thanks to anonymous reviewers are not appropriate.

Conflict of Interest Statement

Authors will be asked to provide a conflict of interest statement during the submission process. For details on what to include in this section, see the section 'Conflict of Interest' in the Editorial Policies and Ethical Considerations section below. Submitting authors should ensure they liaise with all co-authors to confirm agreement with the final statement.

Main Text File

As papers are double-blind peer reviewed, the main text file should not include any information that might identify the authors.

The main text file should be presented in the following order:

i. Title, abstract, highlights and key words;
ii. Main text;
iii. References;
iv. Tables (each table complete with title and footnotes);
v. Figure legends;
vi. Appendices (if relevant).

Figures and supporting information should be supplied as separate files.

Abstract

All manuscripts should contain an abstract of up to 200 words. An **abstract** is a concise summary of the whole paper, not just the conclusions, and is understandable without reference to the rest of the paper. It should contain no citation to other published work. It must be structured, under the sub-headings: Objective; Method; Results; Conclusions.

Highlights

Highlights are mandatory for European Eating Disorders Review. These should appear as three bullet points that convey the core findings of the article.

Keywords

Include up to five **keywords** that describe your paper for indexing purposes.

References

References should be prepared according to the Publication Manual of the American Psychological Association (6th edition). This means in text citations should follow the author-date method whereby the author's last name and the year of publication for the source should appear in the text, for example, (Jones, 1998). The complete reference list should appear alphabetically by name at the end of the paper. A sample of the most common entries in reference lists appears below. Please note that a DOI should be provided for all references where available. For more information about APA referencing style, please refer to the APA FAQ. Please note that for journal articles, issue numbers are not included unless each issue in the volume begins with page one.

Tables

Tables should be self-contained and complement, not duplicate, information contained in the text. They should be supplied as editable files, not pasted as images. Legends should be concise but comprehensive – the table, legend, and footnotes must be understandable without reference to the text. All abbreviations must be defined in footnotes. Footnote symbols: †, ‡, §, ¶, should be used (in that order) and *, **, *** should be reserved for P-values. Statistical measures such as SD or SEM should be identified in the headings.

Figure Legends

Legends should be concise but comprehensive – the figure and its legend must be understandable without reference to the text. Include definitions of any symbols used and define/explain all abbreviations and units of measurement.

Figures

Although authors are encouraged to send the highest-quality figures possible, for peer-review purposes, a wide variety of formats, sizes, and resolutions are accepted. Click here for the basic figure requirements for figures submitted with manuscripts for initial peer review, as well as the more detailed post-acceptance figure requirements.

Additional Files

Appendices

Appendices will be published after the references. For submission they should be supplied as separate files but referred to in the text.

Supporting Information

Supporting information is information that is not essential to the article, but provides greater depth and background. It is hosted online and appears without editing or typesetting. It may include tables, figures, videos, datasets, etc. Click here for Wiley's FAQs on supporting information.

Note: if data, scripts, or other artefacts used to generate the analyses presented in the paper are available via a publicly available data repository, authors should include a reference to the location of the material within their paper.

If a manuscript describes a new approach and/or technological approach, authors are encouraged to include a small demo video – no more than 60 seconds long.

General Style Points

The following points provide general advice on formatting and style.

- **Language:** The language of the journal is English.
- **Abbreviations:** In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Initially, use the word in full, followed by the abbreviation in parentheses. Thereafter use the abbreviation only. Distinction should be made between capital and lower case letters, between the letter O and zero, between the letter I and number one and prime, between k and kappa.
- **Units of measurement:** Measurements should be given in SI or SI-derived units. Visit the Bureau International des Poids et Mesures (BIPM) website at www.bipm.fr for more information about SI units.
- **Numbers:** numbers under 10 are spelt out, except for: measurements with a unit (8mmol/l); age (6 weeks old), or lists with other numbers (11 dogs, 9 cats, 4 gerbils).
- **Trade Names:** Chemical substances should be referred to by the generic name only. Trade names should not be used. Drugs should be referred to by their generic names. If proprietary drugs have been used in the study, refer to these by their generic name, mentioning the proprietary name and the name and location of the manufacturer in parentheses.

Wiley Author Resources

Manuscript Preparation Tips: Wiley has a range of resources for authors preparing manuscripts for submission available here. In particular, authors may benefit from referring to Wiley's best practice tips on Writing for Search Engine Optimization.

Editing, Translation, and Formatting Support: Wiley Editing Services can greatly improve the chances of a manuscript being accepted. Offering expert help in English language editing, translation, manuscript formatting, and figure preparation, Wiley Editing Services ensures that the manuscript is ready for submission.

SCHOOL OF PSYCHOLOGY

DOCTORATE IN CLINICAL PSYCHOLOGY

EMPIRICAL PAPER

Emotion Recognition and Set-Shifting in Women with Anorexia Nervosa

Trainee Name	**Royston Hall**
Primary Research Supervisor (50%)	**Dr Nick Moberly** Senior Lecturer, School of Psychology, University of Exeter
Secondary Research Supervisor (50%)	**Dr Ian Frampton** Senior Lecturer, School of Psychology, University of Exeter, Clinical Psychologist, Eating Disorders Service, Cornwall NHS Foundation Trust
Target Journal	European Eating Disorders Review.
Word Count	7704 (not including abstract, table of contents, references, footnotes, appendices)

Submitted in partial fulfilment of requirements for the Doctorate Degree in Clinical Psychology, University of Exeter

Signature: 

Abstract

Objective: Neuropsychology models of anorexia nervosa (AN) propose that cognitive difficulties including poor Emotion Recognition (ER) and set-shifting ability may be central to the development and maintenance of eating pathology. This study aimed to test the central positions of such models by assessing specific ER difficulties in AN as well as the relationship between ER deficits and set-shifting performance.

Methods: Fifty-one women were assessed (25 with AN; *M* = 28.20 *SD* = 8.69 and 26 control *M* = 21.27 *SD* = 5.10) on a novel measure of ER, a set-shifting test and self-report questionnaires concerning co-morbid factors.

Results: The data did not reveal a global difference in ER or set-shifting performance between groups. Specific hypotheses of ER deficits in AN were also not met as performance on individual emotions was comparable between groups. There was an unexpected negative correlation between disgust recognition and set-shifting performance, however, this was only significant across the whole sample. ER performance was not related with any confounding factors.

Conclusions: Despite an abundance of research supporting the position of social cognitive difficulties in AN, the current study failed to find global or specific deficits in ER in the present sample. Similarly, ER performance was not related to set-shifting as proposed by neuropsychological models of AN aetiology. Possible explanations for a lack of difference observed using this novel ER task are explored and future directions for evaluating ER in AN are discussed.

Keywords*: Anorexia Nervosa, Emotion Recognition, Set-Shifting, Neuropsychology, Social Cognition*

Introduction

Anorexia nervosa (AN) is a complex condition of unclear aetiology that is characterised by nutritional restriction relevant to bodily requirements and cognitive distortions particularly relating to body image (American Psychological Association, 2013). This mental health condition, prevalent in less than 1% of women and even fewer men, has one of the highest rates of mortality and chronicity of any mental health condition (Arcelus, Mitchell, Wales, & Nielsen, 2011). Many factors have been indicated in the development and maintenance of AN including neurobiological and genetic vulnerabilities (Kaye, Wierenga, Bailer, Simmons, & Bischoff-Grethe 2013), cognitive functioning difficulties (Rose, Davis, Frampton, & Lask, 2011) and social and emotional difficulties (Russell, Schmidt, Doherty, Young, & Tchanturia, 2009) among many others. Despite the diversity of research, few studies have considered these biological, cognitive and socioemotional factors in combination (Culbert, Racine, & Klump, 2015).

Conceptual neuropsychology models that encompass all these factors have been developed such as that of Southgate, Tchanturia, and Treasure (2005). This model of AN is presented in simplified form in Figure 1. What differentiates this model from other models of AN is the prediction that difficulties in social and emotional information processing may perpetuate pathology and be underpinned by neurodevelopment secondary to both genetic factors and environmental experience (Brockmeyer et al., 2016). The cognitive interpersonal model builds on this idea to emphasise how neurocognitive abnormalities might impact social and emotional functioning that can perpetuate eating pathology in AN (Treasure & Schmidt, 2013).

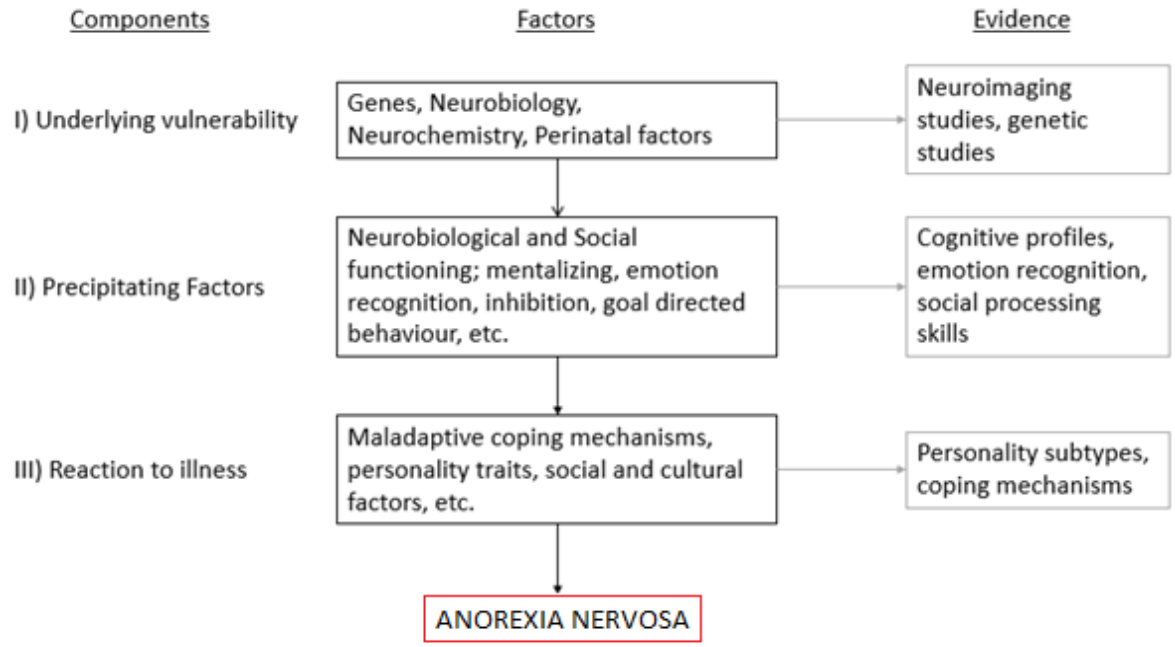

Figure 1. A simplified biopsychosocial model of AN aetiology (based on Southgate et al., 2005).

Emotion Recognition

Neurobiological abnormalities such as functional differences in insula activity (Kaye et al., 2013) as well as neurocognitive differences such as executive dysfunction (Zakzanis, Campbell, & Polsinelli, 2010) are well reported in AN. However, social cognition difficulties are more contentious. Social cognition encompasses a broad range of skills (for a review see Oschner, 2008), but it is emotion recognition in faces (ER) that has received the most attention in the AN literature. ER can be defined as out ability to correctly identify the emotional state of others, typically through facial expressions. As understanding the emotional experience of others is a pre-requisite to understanding our own emotional experience, difficulty with ER may create a vulnerability towards maladaptive emotion regulation strategies (Brockmeyer et al., 2014). Indeed, some authors

suggest that eating pathology can be explained in part as a difficulty with emotion regulation that contributes towards the development and maintenance of maladaptive regulatory strategies such as bingeing and restriction (Haynos & Fruzzetti, 2011), Thus, it is possible that difficulties recognising emotions in others may create vulnerability towards poor emotion regulation skills and subsequently eating pathology. Indeed, many studies support the position that individuals with AN have a deficit recognising emotion (Jänsch, Harmer, & Cooper, 2009; Kanakam, Krug, Raoult, Collier, & Treasure, 2013). However, other authors have failed to find a significant difference between AN and controls (Brewer, Cook, Cardi, Treasure, & Bird, 2015; Medina-Pradas, Navarro, Alvarez-Moya, Grau, & Oliols, 2012).

Despite the abundance of literature and a rich theoretical basis for exploring ER in AN, several methodological limitations in ER studies make robust conclusions challenging. ER Tasks often use static face stimuli (Kucharska-Petura, Nikolaou, Masiak, & Treasure, 2002) or only display images of eyes (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Russell et al., 2009) and as such are not representative of real-world stimuli. Equally, few studies have explored different levels of emotional intensity (Rosenberg, McDonald, Dethier, Kessels, & Westbrook, 2014), or the impact of attentional biases on emotion selection (Cardi, Matteo, Corfield, & Treasure, 2013). Other task variations including stimulus presentation time, number of emotions displayed and response method (e.g. forced choice or free recall) may also influence the results of ER tasks in the literature (Harrison Sullivan, & Tchanturia, 2009). Finally, much of the previous research has not hypothesised or examined performance on specific emotions such as disgust and anger that may be pertinent to understanding this complex disorder.

Disgust and Anger

Disgust sensitivity has been found to be associated with eating disorder psychopathology as well as a moderator of the relationship between disordered eating and suicidal ideation (Chu, Bodell, Ribiero, & Joiner, 2015; Moncrieff-Boyd, Bryne, & Nunn, 2013). Neurobiological hypotheses such as the dysfunctional insula hypothesis offer explanations for this phenomenon (Nunn, Frampton, Gordon, & Lask, 2008). The insula is a brain region that is involved in integrating information across distributed networks and is implicated in the processing of introspective awareness as well as disgust recognition and our ability to shift between emotional states (Gasquoine, 2014; Simmons, Strigo, Matthews, Paulus, & Stein, 2009; Simmons et al., 2012). The insula hypothesis proposes that structural and functional abnormalities in the insula are associated with specific cognitive and affective impairment that perpetuate the development and maintenance of AN (Nunn, Frampton, Fuglset, Törzsök-Sonnevend, & Lask, 2011). The insula model sits comfortably with neuropsychological models of AN such as that of Southgate et al. (2005) by proposing that neurobiological differences may impair both executive functioning and social cognition ability in this population. The core prediction of the insula model, that individuals with AN will demonstrate reduced ability to recognise disgust, has not yet been explicitly examined in ER tasks.

Studies have also shown that social stimuli are interpreted as appearing angrier and more rejecting in AN compared to controls (Cardi, et al., 2013; Gilon et al., 2018; Harrison, Tchanturia, & Treasure, 2010). Of interest is the tendency for disgust to be interpreted as anger when it appears more ambiguous, a trend that was found to be a predictive factor for the vulnerability of eating pathology in a general population study (Jänsch, Harmer, & Cooper, 2009; Pringle, Harmer, &

Cooper, 2010). Fox et al. (2013) further explored this phenomenon and presented an experiment in which anger inductions resulted in participants with AN reporting greater internal disgust. The authors thus concluded that there might be a coupling of anger and disgust in AN. The neural underpinning of emotional experience suggests that similar pathways are active both when observing emotional experience in others and when experiencing the same emotion ourselves (Bastiaansen, Thioux, & Keysers, 2009; Wicker et al., 2003). As such, individuals with AN who mislabel their own internal emotional experience may also mislabel the experience of others similarly, perpetuating social difficulties. Thus, if individuals with AN are misinterpreting emotions as anger, it is important to examine if there is a bias towards recognising anger in ER tasks.

Emotion Recognition and Cognitive Ability

Many factors may impact upon an individual's ability to identify emotion in others correctly but one area that has been largely overlooked is set-shifting ability. This executive function is defined by our ability to switch between different tasks in response to changing goals, for example, the ability to try new strategies when old strategies are no longer effective. Set-shifting has been reported to be a robust deficit in AN (Danner et al., 2012), however, due to the heterogeneity of executive functioning measures, recent research suggests it may only be present in a subset of the AN population (Rose et al., 2016). While relationships between ER and set shifting have not been explored in AN, they are well reported in other clinical conditions in which reduced set-shifting ability is common such as schizophrenia (Lee, Lee, Kweon, Lee, & Lee, 2009) and depression (Uekermann, Abdel-Hamid, Lehmkaemper, Vollmoeller, & Daum, 2008). Importantly, these associations are only

present in clinical populations but not in healthy controls. Southgate et al (2005) hypothesise that associations between reduced executive and emotional functioning in AN may exist due to an interruption in the emergence of collaborative brain functioning in adolescence (Luna, 2009). Therefore, examining the relationship between set-shifting and ER in a population with and without AN will test the central proposition of this neuropsychology model. If correct, it is anticipated that impaired executive difficulties contribute towards reduced social cognitive ability in AN but not in controls.

Confounding Variables

Comorbidities exist in over 55% of individuals with AN (Swanson, Crow, Le Grange, Swendsen, & Merikangas, 2011), many of which have been found to be associated with impaired emotion processing (Surcinelli et al., 2006; Surgladze, Young, Senior, Brébion, & Travis, 2004). While several studies have explored the relationship of co-morbid factors such as alexithymia on ER (Brewer et al., 2015; Courty, Godart, Lallan, & Berthoz, 2015; Dapelo, Surguladze, Morris, & Tchanturia, 2016), their relative contributions remain unclear, and further evaluation is warranted. Similarly, there is some evidence that confounding symptoms such as depression may contribute towards set-shifting difficulties in AN (Giel et al., 2012). Finally, illness factors in AN such as BMI and degree of self-reported eating pathology symptoms need to be considered to determine if ER difficulties are a trait of ED (Harrison, Tchanturia, & Treasure, 2010) or are due to neurophysiological changes associated with starvation (McCormick et al., 2008).

Aims and Hypotheses

The present study aimed to further explore ER difficulties in AN through a measure that is novel to this literature base and hopes to overcome some of the limitations used in previous tasks. Based on neurobiological hypotheses such as the insula model, specific predictions concerning disgust and anger recognition in AN were tested. Factors that may be prerequisites for ER ability as predicted by neuropsychological models of AN were also examined through a validated measure of set-shifting. Finally, to consider the degree of influence confounding factors may have on any findings reported, a series of clinical questionnaires were also completed.

Emotion Recognition Performance in Anorexia Nervosa

Hypothesis 1a: With a novel measure, it will be possible to identify ER differences in AN, specifically that individuals with AN will correctly identify fewer emotions compared to HCs.

Hypothesis 1b: *In line with the insula model of AN it is also hypothesised that individuals with AN will correctly identify fewer faces depicting disgust than controls and misinterpret faces as angry as demonstrated by a bias towards identifying faces as angry*.

The Relationship Between Set-Shifting and Emotion Recognition ability

Hypothesis 2a: In line with neuropsychological models of AN, global or specific (e.g. disgust) ER performance will correlate with performance on measures of set-shifting in the AN group but not in the control group.

Hypothesis 2b: Any relationship between ER and set-shifting will remain robust when accounting for confounding factors such as social anxiety, general anxiety, depression, alexithymia and BMI in the AN group.

Methods and Materials

Design

The study used a cross-sectional experimental design (manipulation of facial emotion expression) with between-subjects comparisons looking at the difference between current patients with AN and non-clinical controls on core dependent variables: ER and set-shifting ability.

Participants

Due to the higher prevalence of females diagnosed with AN and recognised differences in ER across genders (Hoffman, Kessler, Eppel, Rukavina, & Traue, 2010), a female sample was recruited. Similarly, to remove the influence of neurodevelopment (Blakemore & Choudhury, 2006; Horning, Cornwell, & Davis, 2012), only individuals over the age of 18 were invited to participate. Twenty-five participants with a current diagnosis of AN were recruited from inpatient units and outpatient services across the South West of England. All clinical participants were recruited through advertisements in waiting rooms and presentations at inpatient meetings. All clinical participants met DSM-5 criteria for AN as confirmed by clinical lead collaborators at each recruitment site. Weight restored participants were included in the sample as previous research suggests set-shifting and ER are affected independent of weight status in AN (Danner et al., 2012; Harrison et al., 2010). Three participants in the AN group had a BMI higher than 18.5 kg/m^2 though

each maintained clinically significant scores on a measure of eating pathology. BMI in the AN group ranged from 11.7 kg/m^2 to 23.8 kg/m^2 with an average of 16.14 kg/m^2.

Twenty-six controls were recruited from the University of Exeter's Psychology Research Participation System. Control participants had no current or previous history of disordered eating and did not meet the clinical cut off on the eating disorder examination questionnaire (EDE-Q) determined as an average score of 4 or above across all items. All control participants had an EDE-Q score of <2. For all participants, comorbidities including anxiety and depression were not reasons for exclusion; however, participants presenting with other axis I or II disorders or a history of head injury were excluded in the present study. All participants were asked to disclose any medication use. Two clinical participants disclosed use of antidepressant medication and were kept in the final sample. One potential participant was excluded due to multiple medication use that impacted significantly upon fatigue. All participants were offered the opportunity to enter a raffle for shopping vouchers upon completion and control participants also received course credits as part of the University Research Participation System.

Power Analysis

G*Power was used to calculate the minimum sample sizes needed based on effect sizes obtained in previous studies (Faul, Erdfelder, Lang, & Buchner, 2007). For all analyses reported here, alpha was set to .05 and power at .8. A large effect size was assumed for hypothesis one based on Cohen's criteria for comparison of independent means (Cohen, 1992). The large effect size was determined by examination of previous research using the same ER test to compare groups which

found effect sizes of *d* = .9 with sixteen participants (Mullen-Glodde, 2015). For hypothesis 1a, a mixed ANOVA would require a total sample size of at least 76 participants assuming a large effect size (*f* = .40). Subsequent independent samples t-tests would require a minimum of 42 participants (21 in each condition) to address hypothesis 1b. To reliably detect a medium effect size of up to 15% of the variance in hypothesis 2a, a hierarchical regression between ER and set-shifting with the group as a moderating variable would require a minimum of 68 participants (34 in each condition) (Miles & Shevlin, 2001). Finally, for hypothesis 2b, determining whether a relationship between ER and set-shifting remains robust after partialling out covariates using further hierarchical regression analysis, 78 participants would be required across the whole study given an average effect size of *f* = .15. The final recruited sample had sufficient power to explore hypothesis 1b but may be underpowered for the remaining hypotheses.

Behavioural Measures

The Emotion Recognition Task (ERT; Cambridge Cognition Ltd) is a ten-minute computerised task in which participants are briefly presented with 354 x 464-pixel photographs of facial stimuli. Participants are subsequently required to indicate the emotion depicted in a forced choice response (happy, sad, angry, disgusted, surprised, and fearful, see Figure 2a). The faces presented are from a Caucasian male and female, and each of the six emotions are represented along eight levels of intensity (see Figure 2b) using a computer-based morph sequence (Bamford et al., 2015). The 100% intensity stimuli were generated by collating photographs of the same emotion expression, and a prototypical face was created using a neutral expression combined with equal proportions of each of the six emotions to avoid the

neutral face being perceived as threatening (Skinner & Benton, 2010; Yoon, & Zinbarg, 2007). A total of 96 stimuli were used (two sexes, six emotions, eight levels of intensity) each of which was presented only once in a randomised order.

The ERT can be considered to have greater ecological validity than previous measures due to the ambiguity of emotion in faces presented which is more reflective of social interaction (Jhung et al., 2010). This task has recently become part of the COTNAB assessment battery which is commercially available, and research into psychometric properties such as validity and reliability are ongoing. The task and output of data was managed with E-prime software (PST Inc, Sharpsburg, PA, USA). Data from the ER task was analysed by collecting an unbiased hit rate (UHR) which was calculated based on both the number of hits and false alarms for each emotion (Wagner, 1993). Also, total hit rate (THR) and total false alarm rate (FAR) scores were collected across all trials and for each emotion as these are the measures most frequently reported in the literature that allow for direct comparison with previous studies.

At the start of each trial, a fixation cross appears in the centre of the screen and is replaced with a face stimulus for 500ms. Following stimulus presentation, a noise mask is briefly shown to avoid afterimage effects. Participants are then presented with a selection screen with six emotion words representing the six basic emotions measured in the task (see Figure 3). Participants are encouraged to respond as quickly as possible but do not have an upper time limit to respond. Once participants have selected a response, the next trial automatically begins.

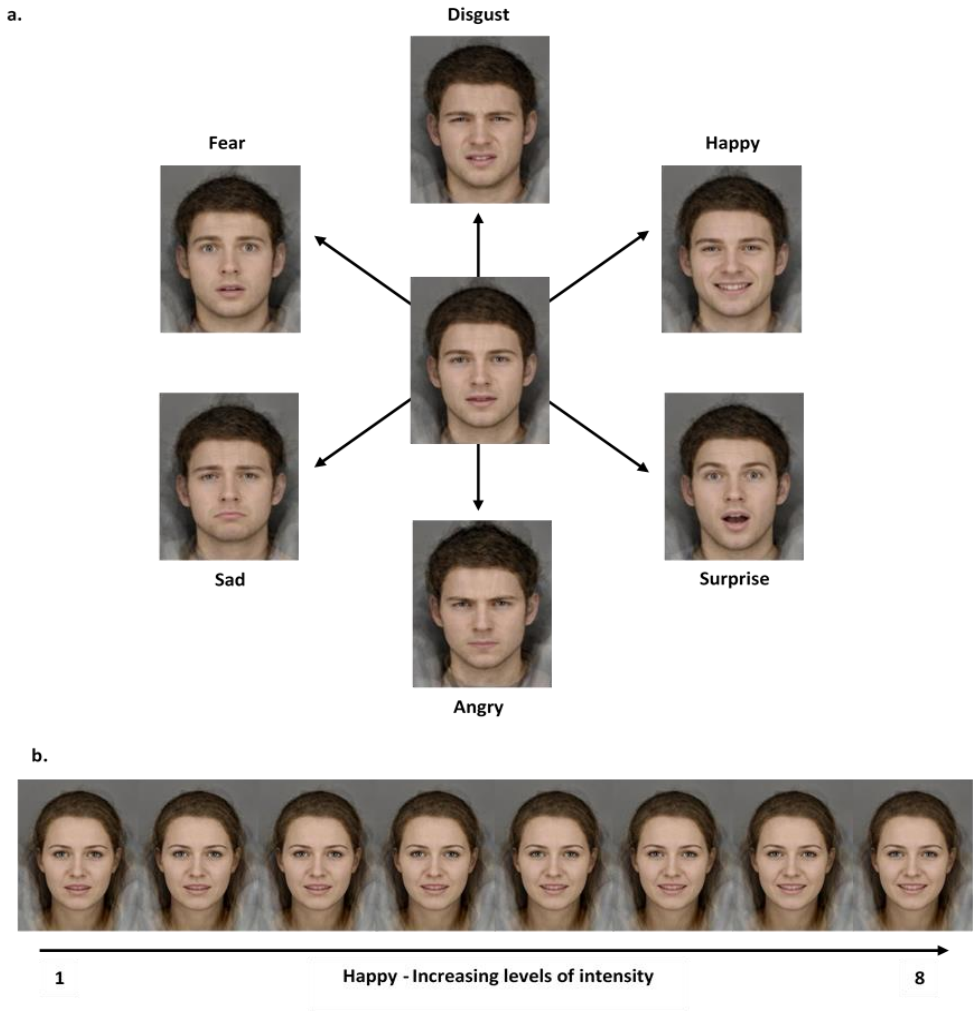

Figure 2 - Examples face stimuli used in the ERT.
a. Prototypical and the most intense level of each of the emotions.
b. Eight levels of intensity used to represent happiness.

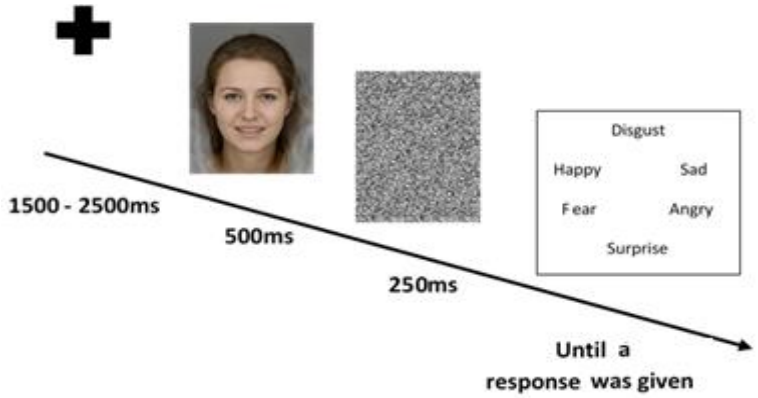

Figure 3 – Procedure for each trial of the ERT

The Penn Conditional Exclusion Test (PCET; Kurtz, Ragland, Moberg, & Gur, 2004) is a measure of set-shifting ability taken from the PENN-State Neuropsychological Testing Battery. This task measures abstraction, concept formation and cognitive flexibility with a design like the Wisconsin Card Sorting Test (Kurtz, Wexler, & Bell, 2004). The task uses the "Odd Man Out" principle in which participants are required to identify which object does not belong in an array of four based on one specific principle (e.g., size, shape or line thickness). Participants are not told the correct principle on any trial and must be guided by feedback on their accuracy. The principle changes without warning when a participant achieves ten consecutive correct answers. Participants have 48 trials to get ten consecutive answers correct in each principle; otherwise, the task is terminated. A total of 24 trials are presented in a non-random order until the participant completes each sorting principle, or the task is terminated. The test score is based on the number of perseverative errors made (e.g., errors made when participants continue to follow an incorrect principle). This task is part of a larger battery that has established reliability and validity for research purposes (Gur et al., 2010).

Figure 4 – Example of a trial of the PCET

Clinical Measures

The Eating Disorders Examination-Q (EDE-Q) is a 41-item self-report questionnaire designed to assess cognitive and behavioural features of eating disorders across four subscales: restraint, eating concern, shape concern and weight concern (Fairburn & Beglin, 1994). Each question on this measure ranged from a score of zero to six and demonstrated good internal consistency (Cronbach's alpha ranging from .89-.97 for different subscales). This scale has good reliability and validity reported in the literature (Mond, Hay, Rodgers, Owen, & Beumont, 2004).

The Brief Fear of Negative Evaluation-Straightforward items (BFNE-S) is an eight-item self-report questionnaire measuring concern around social judgement based on the 12-item BFNE (Carleton, Collimore, & Asmundson, 2007). Each item is rated on a 5-point Likert scale ranging from one to five. Evaluation fear is thought to be a core component of social anxiety and has theoretical links to sensitivity to punishment which is often reported to be exaggerated in AN (Jappe et al., 2011). The scale demonstrated excellent internal consistency (Cronbach's alpha = .85) and average inter-item correlation of .67.

Patient Health Questionnaire-9 (PHQ-9) was used to measure depressive symptoms (Kroenke, Spitzer, & Williams, 2001). This measure has nine self-report items with a four-point Likert scale ranging from zero to three. It is well recognised in clinical practice as well as research and has demonstrated excellent sensitivity (.93) and specificity (.85) for detecting depression (Wittkampf et al., 2009). This scale also demonstrated high reliability in the present sample (Cronbach's alpha = .94), consistent with previous studies (Sherratt & MacLeod, 2013).

Generalized Anxiety Disorder-7 (GAD-7) was also employed to measure anxiety symptoms (Spitzer, Kroenke, Williams, & Lowe, 2006). Using the same Likert

scale as the PHQ-9, this seven-item self-report tool is also regularly applied in clinical and research practice. Consistent with previous reports, the present sample reported high reliability (alpha = .93) of GAD7 items (Sherratt & MacLeod, 2013).

Toronto Alexithymia Scale (TAS) was used to measure alexithymia, this 20 item self-report questionnaire asks participants to rate themselves on a five-point Likert scale ranging from one to five (Bagby, Parker, & Taylor, 1994). This task demonstrated good internal consistency (Cronbach alpha =.86) consistent with reliability reported in the literature (Parker, Taylor, & Bagby, 2003).

Procedure

Clinical participants were identified and screened for suitability by the local collaborator at each site. The testing procedure took between 45-60 minutes and was completed at NHS eating disorder clinic sites. Non-clinical participants were recruited through the University's online booking system and tested in experimental testing rooms at the University. NHS ethical approval was granted by the London Queen Square Health Research Authority, and the University of Exeter Ethics Committee (See Appendix A-B). After gaining informed consent, participants completed the PCET and the ERT on a University research laptop followed by paper clinical measures: PHQ-9, GAD-7, EDE-Q, BFNE-II, TAS.

Analytic Plan

Inspection of the data revealed that all clinical measures apart from BMI were not normally distributed across the entire sample; group mean rank differences were therefore assessed with non-parametric Mann-Whitney tests.

Unbiased hit rate (UHR) for ER was normally distributed, and differences in these scores across groups were explored using a 2 (group) x 6 (emotion) mixed ANOVA. There were no UHR outliers in the data, as assessed by inspection of a boxplot for values greater than three box-lengths from the edge of the box. There was homogeneity of variances, as assessed by Levene's test ($p > .05$) and a Box's test of equality of covariance matrices confirmed homogeneity of covariances ($p = .41$). Mauchly's test of sphericity indicated that the assumption of sphericity was violated for the two-way interaction, so corrected Greenhouse Geisser F-values were reported.

Overall total hit rate (THR) and false alarm rate (FAR) scores across all trials were normally distributed and were analysed through an independent samples t-test. There were no outliers in the data, as assessed by inspection of boxplot graphs for values greater than three box-lengths from the edge of the box. THR and FAR for disgust and anger were not normally distributed as assessed by Shapiro-Wilk's test ($p < .05$) so Mann-Whitney tests were used to compare mean rank differences between groups.

Set shifting was calculated by total perseverative errors on the PENN Conditional Exclusion Task (PCET). One case from the control group was removed from the following analysis as they failed to complete the first set of the set-shifting task. Total perseverative error scores were not normally distributed, so the data was transformed using consecutive log10 transformations. Bivariate correlations were conducted to determine whether there was an association between ER performance and set-shifting in each group. To further examine whether these associations differed by group, a moderation analysis was conducted using multiple regression with ER performance as the outcome, (mean-centred) perseverative errors as the

predictor and group as the moderator (Frazier, Tix, & Barron, 2004). Assumptions of linearity and normality of residuals were upheld, and there was no evidence of multicollinearity. The relative contributions of EDE-Q, BMI, depression, anxiety and alexithymia scores on observed emotion recognition or set-shifting deficits were assessed using appropriate multiple regression analyses to determine a model of best fit.

Results

Participant characteristics

Table 1 presents descriptive statistics for demographic and clinical characteristics in the AN and HC groups. As expected, the AN group scored significantly higher on all clinical scales. Similarly, the AN group had a significantly lower BMI than the HC group. There was no significant group difference in years of education; however, the mean age of the AN group was greater than the control group.

Emotion Recognition Differences

To address hypothesis 1a, a 2 (group) x 6 (emotion) mixed ANOVA revealed that there was no statistically significant group difference in UHR, $F(1, 49) = .76$, $p = .68$, partial $\eta^2 = .004$. However, there was a statistically significant difference across emotions in UHR, $F(3.26, 159.80) = 15.01$, $p < .0005$, partial $\eta^2 = .24$. Bonferroni-corrected pairwise comparisons revealed that, across the entire sample, participants were significantly poorer at recognising fear and surprise than other emotions, all p values < .028 (see Figure 5). There was no statistically significant interaction between group and emotion on UHR, $F(3.26, 159.80)$

= .44, p = .74, partial η^2 = .01, suggesting that the pattern of recognition across emotional expressions did not differ significantly between groups, partially addressing hypothesis 1b. Further, an independent samples t-test revealed that there was no significant difference in overall THR scores between the AN group and controls, $t(49) = .72$, $p = .47$, $d = .20$.

Table 1.
Comparison of demographic and clinical characteristics between anorexia nervosa patients and controls.

Characteristics	Anorexia Nervosa (*n = 25*)		Control Group (*n = 26*)		Test Statistic *U*	*p*	*d*
	Mean rank	*Range*	*Mean Rank*	*Range*			
Age	34.22	19-52	18.10	18-37	119.50	<.0005	1.32
Education (years)	28.46	9-16	23.63	11-18	263.50	.22	.35
EDE-Q: Restraint	35.80	0-4.4	16.58	0-0.8	80.00	<.0005	1.71
EDE-Q: Eating	38.44	0.4-5.8	14.04	0-0.8	26.50	<.0005	3.06
EDE-Q: Shape	37.94	1-6	14.52	0.1-4.1	26.50	<.0005	2.61
EDE-Q: Weight	37.50	1.2-6	14.94	0-2.6	37.50	<.0005	2.35
EDE-Q: Total	37.80	0.6-4.4	14.56	0-2	30.00	<.0005	2.05
PHQ9	38.30	7-27	14.17	0-14	17.50	<.0005	2.85
GAD7	37.02	2-21	15.40	0-18	49.50	<.0005	2.17
BFNE-II	36.20	32-60	16.19	15-60	70.00	<.0005	1.47
TAS-20	37.52	39-78	14.92	31-60	37.00	<.0005	2.40
BMI	Mean = 16.14 ± 2.79	11.7-23.8	Mean = 20.78 ± 2.25	16.6-25.1	$t = 6.56$	<.0005	1.83

Note. BFNE = Brief Fear of Negative Evaluation; BMI = Body Mass Index; EDE-Q = Eating Disorder Examination – Questionnaire; GAD-7 = Generalised Anxiety Disorder Questionnaire; PHQ-9 = Physical Health Questionnaire; TAS-20 = Toronto Alexithymia Scale

To further test hypothesis 1b, Mann-Whitney tests were run to explore differences in THR and FAR in detecting disgust and anger. THR for disgust did not differ by group, $U(49) = 320.5$, $p = .93$, $d = .02$, however, there was a trend for a lower THR of anger in the AN group which approached significance, $U(49) = 226$, $p = .061$, $d = .07$. FAR rates of disgust and anger did not significantly differ by group, $U(49) = 322$, $p = .96$, $d = .02$, and $U(49) = 321$, $p = .94$, $d = .02$ respectively.

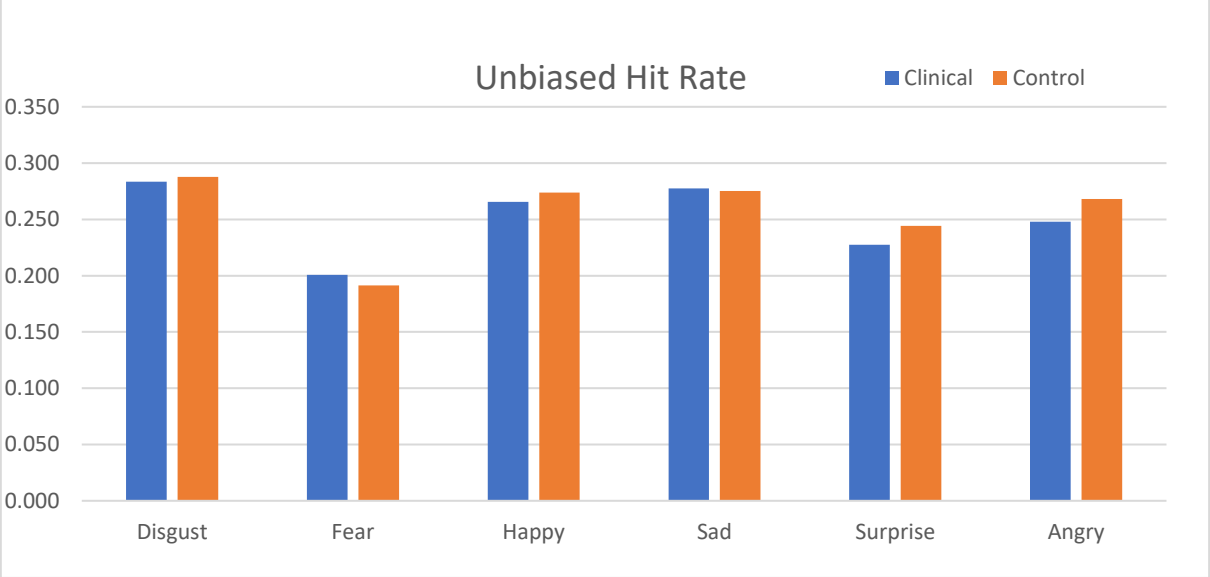

Figure 5 – Mean unbiased hit rate across each emotion for anorexia nervosa and control group.

To explore hypothesis 2a, an independent samples t-test confirmed that there was no significant difference in perseverative errors between the clinical and control groups, $t(48) = -.07$, $p = .94$, $d = .01$. Pearson's correlation analyses explored the relationship between the transformed perseverative error score and THR across total ER performance as well as UHR for both anger and disgust (see Table 2).

Table 2

Pearson's correlation of emotion recognition ability and perseverative errors.

Emotion	AN group (*n* = 25)		HC group (*n* = 25)		Total Sample (*n* = 50)	
	r	*p*	*r*	*p*	*r*	*p*
THR all emotions	-.16	.45	-.29	.16	-.23	.11
UHR Disgust	-.39	.051	-.39	.06	-.39	.005
UHR Anger	-.21	.32	-.16	.45	-.19	.20

THR = Total hit rate; UHR = Unbiased hit rate

There was no relationship between THR across total emotions within the entire sample or within AN and HC groups, failing to support the hypothesis that there is a relationship between set-shifting and overall ER ability in the AN group. However, a significant relationship was identified between UHR of disgust and perseverative errors across the whole sample, $r = -.39$, $p = .005$, indicating that poorer recognition of disgust correlates with a greater number of perseverative errors on a task of set-shifting. The hypotheses of this study did not predict this finding. This relationship did not quite reach statistical significance in the AN ($r = -.39$, $p = .051$) or HC group ($r = -.39$, $p = .06$) when analysed separately. No other UHR scores for specific emotions correlated with perseverative errors across individual groups or the total sample.

Further, a hierarchical multiple regression was conducted to assess whether the association between perseverative errors were differentially associated with UHR of disgust across the two groups. UHR of disgust was entered as the outcome, perseverative errors as the predictor and group as the moderator, followed by the interaction between perseverative errors and group. Group did not moderate the association between perseverative errors and UHR for disgust, as evidenced by an

increase in total variation explained by the perseverative errors by group interaction of .09%, which was not statistically significant, $F(1, 46) = .009$, $p = .97$.

Finally, hypothesis 2b could not be tested as the relationship between perseverative errors and ER was identified across the whole sample and was not selectively present in the AN group. However, correlations were explored between mood disorder symptoms and self-reported severity of illness (e.g., BMI and EDE-Q scores) and ER performance across the sample (see Table 3). This exploratory analysis revealed that confounding factors did not correlate with ER performance in any group. However, self-reported eating pathology correlated with ER performance in the AN group and self-reported restraint demonstrated the most robust relationship in the AN group and the total sample.

Table 3

Correlational analyses exploring the relationship between confounding factors and ER ability

Confound	AN group (*n* = 25)		HC group (*n* = 26)		Total Sample (*n* = 51)	
	r	*p*	*r*	*p*	*r*	*p*
PHQ9	-.20	.35	-.34	.09	-.24	.09
BFNE-II	.14	.50	-.01	.98	-.021	.88
TAS-20	-.29	.15	-.05	.81	-.18	.21
GAD7	-.29	.15	-.14	.49	-.22	.12
BMI	-.33	.11	.01	.98	-.06	.69
EDE-Q Restraint	-.58	.002	-.07	.74	-.32	.02
EDE-Q Eating Concern	-.45	.03	-.07	.74	-.23	.10
EDE-Q Shape Concern	-.25	.22	.11	.60	-.14	.32
EDE-Q Weight concern	-.28	.17	.11	.60	-.16	.25
EDE-Q Total	-.49	.01	.07	.75	-.22	.12

Note. BFNE = Brief Fear of Negative Evaluation; BMI = Body Mass Index; EDE-Q = Eating Disorder Examination – Questionnaire; GAD-7 = Generalised Anxiety Disorder Questionnaire; PHQ-9 = Physical Health Questionnaire; TAS-20 = Toronto Alexithymia Scale

Additionally, an exploratory hierarchical multiple regression was run to determine if the addition of these factors would significantly influence the strength of any association between ER and set-shifting ability across the whole sample (see Table 4). There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.93. The addition of mood variables (Model 2), led to a non-significant increase in R^2 of .08, $F(4,44) = 1.15$, $p = .35$. Equally, the addition of self-reported illness severity to the prediction of disgust (Model 3) also provided a non-significant increase in R^2 of .011, $F(2, 42) = .29$, $p = .75$. None of the added variables significantly contributed to the variance explained by the basic model and UHR for disgust remained the only significant predictor of perseverative errors in each model.

Table 4

Hierarchical Regression Analysis examining the strength of relationship between perseverative errors, UHR for disgust and confounding factors

	B	SE B	Beta	*p*	Variance Explained
Model 1					
Constant	1.50	.53			
UHR Disgust	-2.64	.90	-.39	.005	R^2 = .15** R^2 adjusted .13
Model 2					
Constant	1.08	.88			
UHR Disgust	-2.51	.90	-.37	.008	
PHQ9	-.00	.03	-.03	.92	
BFNE-II	.03	.02	.40	.06	
TAS-20	-.07	.02	-.22	.36	
GAD7	-.00	.04	-.02	.95	R^2 = .23* R^2 adjusted = .15
Model 3					
Constant	2.14	1.71			
UHR Disgust	-2.60	.98	-.39	.01	
PHQ9	-.00	.04	-.01	.98	
BFNE-II	.03	.07	.36	.11	
TAS-20	-.07	.02	-.23	.38	
GAD7	-.00	.04	-.03	.94	
BMI	-.04	.06	-.14	.47	
EDE-Q	-.00	.01	-.07	.82	R^2 = .24 R^2 adjusted = .11

Note. * $p<.05$, ** $p<.01$.; BFNE = Brief Fear of Negative Evaluation; BMI = Body Mass Index; EDE-Q = Eating Disorder Examination – Questionnaire; GAD-7 = Generalised Anxiety Disorder Questionnaire; PHQ-9 = Physical Health Questionnaire; TAS-20 = Toronto Alexithymia Scale; UHR = Unbiased Hit Rate

Discussion

The present study aimed to investigate whether global or specific emotion recognition impairments exist in a sample of adult women with AN using a novel and clinically accessible experimental task: the emotion recognition task (ERT). In the context of a neuropsychological model of eating pathology (Southgate et al., 2005), the present study also sought to explore relationships between social information processing skills such as ER and executive functioning ability such as set-shifting. Finally, to address the variability of findings in the ER literature, a series of confounding factors and their association with ER functioning was also considered.

With regards to the first hypothesis, the ERT did not reveal any difference in overall ER performance between the clinical group and controls. While this is contrary to much of the literature in ER and eating disorders (Harrison, Sullivan, et al., 2009; Russell et al., 2009), this is not the first study to fail to demonstrate global emotion recognition difficulties in AN (Brewer et al., 2015; Rothschild-Yakar, Eviatar, Shamia, & Gur, 2011). The lack of findings in the present study may be due to a Type I error as the sample size may be too small to detect meaningful differences between groups. It is also possible that ER deficits in AN are only present in less ambiguous stimuli such that both the clinical and control groups found the morphed stimuli in the present task equally challenging. Indeed, previous studies have reported that ER differences are only present between ED groups and controls when stimuli are less ambiguous (Dapelo, Surguladze, Morris, & Tchanturia, 2017). A further analysis of ER ability at different emotional intensities in the present task may be able to explore this hypothesis further.

In addition to the lack of global difference in ER between groups, comparable performance was observed between groups when recognising faces depicting disgust (Dapelo et al., 2015; Pollatos et al., 2008). While the lack of significance may be due to a lack of power in the present sample, the study also found small effect sizes when exploring differences in disgust recognition between groups. Thus, this finding is contrary to what might be expected by the insula hypothesis that associates reduced processing of both introspective bodily awareness and disgust with functional abnormalities in the insula (Nunn et al., 2008).

It was also hypothesised that there may be a greater false alarm rate for angry faces in the AN group due to the misclassification of disgust faces as anger (Gilon et al., 2018; Jänsch et al., 2009). This hypothesis was not supported and a non-significant trend for the AN group to have a lower THR for emotions depicting anger was observed instead. This trend may be due to social cognitive difficulties such as poor emotion regulation and active suppression and inhibition of negative feelings in AN (Espeset, Gulliksen, Nordbø, Skårderud, & Holte, 2012). Avoidance of negative emotional experience in AN may translate into a reluctance to select this emotional label in the clinical sample, even despite an attentional bias towards negative stimuli (Ioannou & Fox, 2009).

Previous research has reported total hit rate (THR) as discussed above. However, THR data cannot account for attentional biases of participants and few studies have explicitly explored the impact of biases in ER studies (Dapelo, et al., 2016; Jänsch et al., 2009). For example, a participant with a bias towards selecting anger may receive a high THR score for an emotion by selecting this label in most trials. However, we cannot reliably conclude accuracy in distinguishing anger from other emotions as this participant may have selected anger on every trial. The

present study attempted to overcome the limitations of THR data by calculating unbiased hit rate (UHR) from the total number of correct responses as well as the false alarm rates for each emotion (Wagner, 1993). When UHR was analysed in place of THR, the trend for reduced anger recognition in the AN sample was no longer present. The failure to find a difference in ER performance in the present study compared to previous literature may be due to the unbiased analysis reported here. The addition of an unbiased analysis is a strength of the present design that highlights the importance of considering bias when reporting ER in AN, the lack of which may explain some of the variability observed in the ER literature.

With regards to hypothesis two, the present study did not identify any group differences in set-shifting ability as determined by perseverative errors between groups. There was also no correlation between overall ER performance and set-shifting ability in either group, suggesting that there is not a common mechanism underlying performance in these two domains as proposed by the cognitive-interpersonal model (Treasure & Schmidt, 2013) and Southgate et al.'s (2005) model of AN. Interestingly, the present study did find an unexpected relationship wherein greater perseverative errors correlated with reduced UHR for disgust, across the entire sample. Despite plentiful research in other clinical conditions that have found a direct relationship between ER and set-shifting (Lee et al., 2009), these studies do not comment on the relationship between set-shifting and recognition of specific emotions such as disgust. To our knowledge, this is the first study to find such a relationship among both clinical and control participants. While disgust recognition has not been explored with set-shifting, it has been found to be associated with other executive abilities such as inhibition in young adults (Circelli, Clark, & Cronin-Golomb, 2013).

One interpretation of this relationship may relate to inflexibility of eye gaze in response to facial stimuli. Disgust in faces is most easily discerned by observation of the lower half of the face, while most other facial expressions can be reliably identified through the eyes (Calder, Young, Keane, & Dean, 2000). Given the limited presentation time of stimuli, it is advantageous for participants to focus on the top half of the face when trying to determine the emotion presented. However, if participants are inflexible in diverting their eye gaze to the bottom half of the face stimuli and perseverate instead on only the eyes, they are more likely to miss essential cues for recognising disgust. Similarly, older adults with impaired executive function have also been reported to spend longer fixating on the top half of face stimuli than the bottom (Circelli et al., 2013). Future studies exploring this relationship may consider using eye-gaze data to explore this hypothesis further.

The present study sought to explore the association between co-morbid factors and illness severity; BMI, EDE-Q scores, depression, anxiety and alexithymia upon ER. Consistent with previous reports, an exploratory analysis did not reveal a relationship between depression or anxiety symptoms upon ER ability in the clinical or control group (Dapelo et al., 2016). ER performance was not correlated with alexithymia in the AN group, adding further evidence to the position that alexithymia does not underpin ER differences in eating disorder populations (Kessler et al., 2006; Torres et al., 2015).

Finally, contrary to some previous findings, clinical factors such as severity of illness as measured through BMI did not influence ER performance in either group (Oldershaw, Hambrook, Tchanturia, Treasure, & Schmidt, 2010). However, across the whole sample and in the AN group specifically, self-reported restriction demonstrated a significantly stronger relationship with ER ability than other

dimensions of eating pathology including shape and weight concern. Previous studies have suggested that restricting subtypes of AN show greater ER difficulties than bingeing subtypes of AN (Harrison et al., 2010) and there may also be differences in the executive functioning profiles of different AN subtypes (van Autreve, Da Beane, Baeken, van Heeringen, & Vervaet, 2013). In addition, it is also suggested that there are several clusters of cognitive profile present in the AN population (Rose et al., 2016). Such findings highlight the significant heterogeneity of this clinical population such that comparing cognitive performance based on diagnosis alone may be too broad. The lack of finding in the present study and the general variability in the literature may be a consequence of assessing at the level of diagnosis and not at the level of specific cognitive or behavioural phenotypes. Larger neurocognitive studies such as the Ravello profile of AN may wish to include ER tests in their battery to explore whether ER difficulties are associated with specific behavioural or cognitive subtypes of AN (Rose et al., 2011).

The present study has several strengths compared to previous literature exploring ER in eating disorders. For example, the experimental task is a measure that uses morphing of facial stimuli to detect differences in ER ability, improving its ecological validity over previous measures such as the Reading the Mind in the Eyes test (Baron-Cohen et al., 1997; Jhung et al., 2010). The analysis of bias data also allows for a more accurate measure of recognition ability. The ERT is publicly available allowing for easy replication with the potential of clinical application. Finally, the evaluation of a comprehensive set of confounding features that are known to influence ER allow for more robust conclusions of ER performance in AN.

In isolation, the present study does not support the validity of the ERT as a clinical tool for measuring ER in AN. However, that is not to say that understanding

individual differences and difficulties in ER cannot be clinically useful (Penton-Voak, Bate, Lewis, & Munafò, 2012). Indeed, ER intervention studies are already demonstrating meaningful quality of life improvements for individuals with eating disorders (Davies et al., 2012; Money et al., 2011; Tchanturia, Doris, Mountford, & Fleming, 2015). As the current task is available within the public domain for clinician use, it would be beneficial to collect normative data for individuals with AN (Brooks, Sherman, Iverson, Slick, & Strauss, 2011).

Limitations

The present study's modest sample size may have impacted the reliability of the findings reported here. G*Power analysis suggests that with a power of .80 the present sample was able to reliably detect a small effect size (d = 0.3) in a repeated measures ANOVA as per hypotheses 1a and 1b and a medium effect size (d= .66) for a hierarchical regression exploring hypotheses 2a and 2b. Replications using a larger sample size would reduce the risk of type I errors in which a false null hypothesis is incorrectly retained. Analysis using a more extensive clinical sample would improve the power of the present analysis and allow for comparisons between sub-groups of the clinical population (e.g., restricting or bingeing) that were not possible in the present sample.

As well as differences in sample size, sample characteristics may limit the generalisation of the current findings. For example, the current study only recruited female participants as the predominance of individuals who present with this clinical condition are female. There may be differences in emotion processing between men and women (Hoffman et al., 2010) such that ER deficits may be more pronounced in a male sample (Hall & Matsumoto, 2004; Strother, Lemberg, Stanford, & Turberville,

2012). Another limitation is the reliance upon self-report information. Illness severity could only be determined through self-reported pathology as severity is challenging to measure in EDs which are rarely linear, have a poorly identified onset and follow a pattern of remission and relapse (Tierney & Fox, 2009). The use of self-report questionnaires such as the EDE-Q also has limitations as control participants scoring below the clinical threshold may still demonstrate disordered eating. Future studies using a larger sample of control participants may wish to only include participants within the lowest quartile of EDE-Q scores to ensure that there is minimal chance of unrecognised eating pathologies among the control group.

Similarly, clinical participants were not screened based on medication use. It is ecologically valid to include cases with medication use, particularly antidepressants, as these are commonly used in this clinical population. However, a comparison of performance between groups using medication and those not is important as there is some anecdotal evidence that medication can impact ER ability (Jänsch et al., 2009). The clinical population included in the present study had a significantly low BMI. Experimental tasks such as the ER task require sustained attention for up to 15 minutes which may be challenging for severely malnourished participants. As a result, it is possible that confounding factors such as fatigue and attention may have reduced performance on experimental measures. Future studies may wish to employ a formal assessment of attentional capacity to control for this confound. Finally, as the control group was exclusively recruited through undergraduate university students, there was a significant difference in the age of the clinical population and the control group. Both ER and set-shifting performance may be related to changes associated with age and neurological maturation (Horning et al., 2012), potentially impacting the results found in the present sample. Although

previous research has demonstrated that age may not influence ER performance in an AN population (Russell et al., 2008), the difference in age reported between the two groups limits the conclusions that can be drawn here.

While the present study did identify some unusual relationships between disgust recognition and perseverative error scores, data from the PCET task was positively skewed suggesting a ceiling effect in the tasks employed. The PCET was chosen in part due to its accessibility, and short administration time, however, future studies may wish to explore set-shifting with more thorough measures such as the Wisconsin card sorting test (WCST, Puente, 1985). Similarly, the present task looked explicitly at set-shifting, but other executive function difficulties may also relate to ER ability in AN. For example, central coherence is considered a stable deficit in AN as observed through a preference for local over global processing of images (Lang, Lopez, Stahl, Tchanturia, & Treasure, 2014). Difficulties distinguishing different emotional expressions may be impacted upon by ability to either holistically scan faces or scan individual facial features such as furrowed eyebrows (Gery, Mijkovitch, Berthoz, & Soussigan, 2009). A local processing bias may mean that individuals with AN selectively process specific local facial cues making emotions such as anger and disgust harder to differentiate. Further studies may wish to employ a more comprehensive battery of executive functions to explore relationships between neurocognitive ability and social cognition as described in neuropsychological models of AN.

Conclusion

Contrary to the majority of literature in this field, the present study failed to replicate findings of a global or specific emotion recognition difficulty in individuals with AN using a novel task. Comparable performance between clinical and non-clinical groups failed to support neurocognitive models of AN such as the insula hypothesis. Similarly, the present study did not support the social information processing models of AN as there was no relationship between set-shifting ability and ER performance in the AN group. An unusual relationship between disgust recognition and set-shifting was found across the entire sample that may be explained by inflexibility of eye-gaze. Clinically relevant confounding factors did not correlate with ER performance in the present study and did not reduce the significance of the relationship between disgust recognition and set-shifting ability across the entire sample. The failure to find support for the hypotheses in the present study may be due to small sample sizes and sample characteristics. Similarly, assessing cognitive difficulties across a population as heterogeneous as AN is likely to contribute towards the variability in reported ER deficit. Future studies may need to consider moving beyond diagnosis and consider an analysis of ER across behavioural and cognitive subtypes of AN presentation. Equally, a more comprehensive battery of executive tasks may further clarify if neurocognitive difficulties such as executive functioning contribute towards social cognitive difficulties including ER among individuals with AN. Advances in our understanding of these complex abilities will further support the development and targeted implementation of alternative treatment programmes such as cognitive remediation therapy and ER training programmes which have been demonstrated to be beneficial for individuals with AN.

Acknowledgements

I would like to thank Professor Ian Penton-Voak for providing the Emotion Recognition Task used in this study and supporting the interpretation of E-Prime data output. I am also grateful to my supervisors, Dr Nick Moberley and Dr Ian Frampton, for their support and helpful comments on earlier drafts of the manuscript.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders: a meta-analysis of 36 studies. *Archives of General Psychiatry*, 68, 724-731.

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*, 23-32.

Bamford, S., Penton-Voak, I., Pinkney, V., Baldwin, D. S., Munafò, M. R., & Garner, M. (2015). Early effects of duloxetine on emotion recognition in healthy volunteers. *Journal of Psychopharmacology*, *29*, 634-641.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *38*, 813-822.

Bastiaansen, J. A., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 2391-2404.

Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*, 296-312.

Brewer, R., Cook, R., Cardi, V., Treasure, J., & Bird, G. (2015). Emotion recognition deficits in eating disorders are explained by co-occurring alexithymia. *Royal Society Open Science*, *2*, 140382.

Brockmeyer, T., Pellegrino, J., Münch, H., Herzog, W., Dziobek, I., & Friederich, H.-C. (2016). Social cognition in anorexia nervosa: Specific difficulties in decoding emotional but not nonemotional mental states. *International Journal of Eating Disorders*, *49,* 883–890.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 527.

Cardi, V., Matteo, R. D., Corfield, F., & Treasure, J. (2013). Social reward and rejection sensitivity in eating disorders: An investigation of attentional bias and early experiences. *The World Journal of Biological Psychiatry*, *14*, 622-633.

Carleton, R. N., Collimore, K. C., & Asmundson, G. J. G. (2007). Social anxiety and fear of negative evaluation: Construct validity of the BFNE-II. *Journal of Anxiety Disorders, 21*, 131-141.

Chu, C., Bodell, L. P., Ribeiro, J. D., & Joiner, T. E. (2015). Eating Disorder Symptoms and Suicidal Ideation: The Moderating Role of Disgust. *European Eating Disorders Review : The Journal of the Eating Disorders Association*, *23*, 545–552.

Circelli, K. S., Clark, U. S., & Cronin-Golomb, A. (2013). Visual scanning patterns and executive function in relation to facial emotion recognition in aging. *Aging, Neuropsychology, and Cognition, 20*, 148-173.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Courty, A., Godart, N., Lalanne, C., & Berthoz, S. (2015). Alexithymia, a compounding factor for eating and social avoidance symptoms in anorexia nervosa. *Comprehensive Psychiatry*, *56*, 217–228.

Culbert, K. M., Racine, S. E., & Klump, K. L. (2015). Research Review: What we have learned about the causes of eating disorders - a synthesis of sociocultural, psychological, and biological research. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *56*, 1141–1164.

Danner, U. N., Sanders, N., Smeets, P. A., van Meer, F., Adan, R. A., Hoek, H. W., & van Elburg, A. A. (2012). Neuropsychological weaknesses in anorexia nervosa: Set-shifting, central coherence, and decision making in currently ill and recovered women. *International Journal of Eating Disorders*, *45*, 685-694.

Dapelo, M. M., Surguladze, S., Morris, R., & Tchanturia, K. (2016). Emotion recognition in blended facial expressions in women with anorexia nervosa. *European Eating Disorders Review*, *24*, 34-42.

Davies, H., Fox, J., Naumann, U., Treasure, J., Schmidt, U., & Tchanturia, K. (2012). Cognitive remediation and emotion skills training for anorexia nervosa: An observational study using neuropsychological outcomes. *European Eating Disorders Review*, *20*, 211–217.

Espeset, E., Gulliksen, K. S., Nordbø, R. H., Skårderud, F., & Holte, A. (2012). The link between negative emotions and eating disorder behaviour in patients with anorexia nervosa. *European Eating Disorders Review*, *20*, 451-460.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*, 363-370.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Fox, J. R., Smithson, E., Baillie, S., Ferreira, N., Mayr, I., & Power, M. J. (2013). Emotion coupling and regulation in anorexia nervosa. *Clinical Psychology & Psychotherapy*, *20*, 319-333.

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, *51*, 115-134.

Gasquoine, P. G. (2014). Contributions of the insula to cognition and emotion. *Neuropsychology Review*, *24*, 77-87.

Gery, I., Miljkovitch, R., Berthoz, S., & Soussignan, R. (2009). Empathy and recognition of facial expressions of emotion in sex offenders, non-sex offenders and normal controls. *Psychiatry Research*, *165*, 252-262.

Giel, K. E., Wittorf, A., Wolkenstein, L., Klingberg, S., Drimmer, E., Schönenberg, M., ... & Zipfel, S. (2012). Is impaired set-shifting a feature of “pure” anorexia nervosa? Investigating the role of depression in set-shifting ability in anorexia nervosa and unipolar depression. *Psychiatry Research, 200*, 538-543.

Gilon, M. T., Hamdan, S., Bar-Haim, Y., Lazarov, A., Enoch-Levy, A., Dubnov-Raz, G., ... & Stein, D. (2018). Different attention bias patterns in anorexia nervosa restricting and binge/purge types. *European Eating Disorders Review,* 1-9.

Gur, R. C., Richard, J., Hughett, P., Calkins, M. E., Macy, L., Bilker, W. B., ... & Gur, R. E. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *Journal of Neuroscience Methods, 187*, 254-262.

Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion, 4*, 201.

Harrison, A., Sullivan, S., Tchanturia, K., & Treasure, J. (2009). Emotion recognition and regulation in anorexia nervosa. *Clinical Psychology & Psychotherapy, 16*, 348-356.

Harrison, A., Tchanturia, K., & Treasure, J. (2010). Attentional bias, emotion recognition, and emotion regulation in anorexia: state or trait? *Biological Psychiatry*, *68*, 755–61.

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica, 135*, 278-283.

Horning, S. M., Cornwell, R. E., & Davis, H. P. (2012). The recognition of facial expressions: an investigation of the influence of age and cognition. *Aging, Neuropsychology, and Cognition, 19*, 657-676.

Ioannou, K., & Fox, J. R. (2009). Perception of threat from emotions and its role in poor emotional expression within eating pathology. *Clinical Psychology & Psychotherapy*, *16*, 336-347.

Jänsch, C., Harmer, C., & Cooper, M. J. (2009). Emotional processing in women with anorexia nervosa and in healthy volunteers. *Eating Behaviors*, *10*, 184-191.

Jappe, L. M., Frank, G. K. W., Shott, M. E., Rollin, M. D. H., Pryor, T., Hagman, J. O., … Davis, E. (2011). Heightened sensitivity to reward and punishment in anorexia nervosa. *The International Journal of Eating Disorders*, *44*, 317–324.

Jhung, K., Namkoong, K., Kang, J. I., Ha, R. Y., An, S. K., Kim, C. H., & Kim, S. J. (2010). Perception bias of disgust in ambiguous facial expressions in obsessive–compulsive disorder. *Psychiatry Research*, *178*, 126-131.

Jones, L., Harmer, C., Cowen, P., & Cooper, M. (2008). Emotional face processing in women with high and low levels of eating disorder related symptoms. *Eating Behaviors*, *9*, 389-397.

Kanakam, N., Krug, I., Raoult, C., Collier, D., & Treasure, J. (2013). Social and emotional processing as a behavioural endophenotype in eating disorders: a pilot investigation in twins. *European Eating Disorders Review*, *21*, 294-307.

Kaye, W. H., Wierenga, C. E., Bailer, U. F., Simmons, A. N., & Bischoff-Grethe, A. (2013). Nothing tastes as good as skinny feels: the neurobiology of anorexia nervosa. *Trends in Neurosciences*, *36*, 110–120.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606-613.

Kucharska-Pietura, K., Nikolaou, V., Masiak, M., & Treasure, J. (2004). The recognition of emotion in the faces and voice of anorexia nervosa. *International Journal of Eating Disorders, 35*, 42-47.

Kurtz, M. M., Ragland, J. D., Moberg, P. J., & Gur, R. C. (2004). The Penn Conditional Exclusion Test: A new measure of executive-function with alternate forms for repeat administration. *Archives of Clinical Neuropsychology, 19*, 191-201.

Kurtz, M. M., Wexler, B. E., & Bell, M. D. (2004). The Penn Conditional Exclusion Test (PCET): relationship to the Wisconsin Card Sorting Test and work function in patients with schizophrenia. *Schizophrenia Research, 68*, 95-102.

Lang, K., Lopez, C., Stahl, D., Tchanturia, K., & Treasure, J. (2014). Central coherence in eating disorders: an updated systematic review and meta-analysis. *The World Journal of Biological Psychiatry : The Official Journal of the World Federation of Societies of Biological Psychiatry*, *15*, 586–98.

Lee, S. J., Lee, H.-K., Kweon, Y.-S., Lee, C. T., & Lee, K.-U. (2009). The impact of executive function on emotion recognition and emotion experience in patients with schizophrenia. *Psychiatry Investigation*, *6*, 156–62.

Luna, B. (2009). The maturation of cognitive control and the adolescent brain. In Aboitiz, F., & Cosmelli, D. (Ed.), *From attention to goal-directed behavior* (pp. 249-274). Springer, Berlin, Heidelberg.

Medina-Pradas, C., Navarro, J. B., Álvarez-Moya, E. M., Grau, A., & Obiols, J. E. (2012). Emotional theory of mind in eating disorders. *International Journal of Clinical and Health Psychology*, *12*, 189-192.

Miles, J., & Shevlin, M. (2000). *Applying Regression and Correlation: A Guide for Students and Researchers*. SAGE.

Moncrieff-Boyd, J., Byrne, S., & Nunn, K. (2013). Disgust and Anorexia Nervosa: confusion between self and non-self. *Advances in Eating Disorders*, *2*, 4-18.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, *42*, 551-567.

Money, C., Davies, H., & Tchanturia, K. (2011). A case study introducing cognitive remediation & emotion skills training (CREST) for anorexia nervosa in-patient care. *Clinical Case Studies, 10*, 110–121.

Mullen-Glodde, M. (2015). Emotion perception in brain injury patients: Assessing the feasibility of the Bristol Emotion Recognition Task, *Unpublished manuscript.* University of Bristol, UK.

Nunn, K., Frampton, I., Fuglset, T. S., Törzsök-Sonnevend, M., & Lask, B. (2011). Anorexia nervosa and the insula. *Medical Hypotheses*, *76*, 353–357.

Nunn, K., Frampton, I., Gordon, I., & Lask, B. (2008). The fault is not in her parents but in her insula--a neurobiological hypothesis of anorexia nervosa. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *16*, 355–360.

Oldershaw, A., Hambrook, D., Tchanturia, K., Treasure, J., & Schmidt, U. (2010). Emotional theory of mind and emotional awareness in recovered anorexia nervosa patients. *Psychosomatic Medicine, 72*, 73-79.

Ochsner, K. N. (2008). The social-emotional processing stream: five core constructs and their translational potential for schizophrenia and beyond. *Biological Psychiatry*, *64*, 48-61.

Parker, J. D., Taylor, G. J., & Bagby, R. M. (2003). The 20-Item Toronto Alexithymia Scale: III. Reliability and factorial validity in a community population. *Journal of Psychosomatic Research*, *55*, 269-275.

Penton-Voak, I. S., Bate, H., Lewis, G., & Munafò, M. R. (2012). Effects of emotion perception training on mood in undergraduate students: randomised controlled trial. *The British Journal of Psychiatry, 201*, 71-72.

Pollatos, O., Kurz, A. L., Albrecht, J., Schreder, T., Kleemann, A. M., Schöpf, V., ... & Schandry, R. (2008). Reduced perception of bodily signals in anorexia nervosa. *Eating behaviors, 9*, 381-388.

Pringle, A., Harmer, C. J., & Cooper, M. J. (2010). Investigating vulnerability to eating disorders: biases in emotional processing. *Psychological Medicine*, *40*, 645-655.

Puente, A. (1985). Wisconsin card sorting test. *Test Critiques, 4*, 677-682.

Ridout, N., Wallis, D. J., Autwal, Y., & Sellis, J. (2012). The influence of emotional intensity on facial emotion recognition in disordered eating. *Appetite*, *59*, 181-186.

Rose, M., Davis, J., Frampton, I., & Lask, B. (2011). The Ravello Profile: development of a global standard neuropsychological assessment for young people with anorexia nervosa. *Clinical Child Psychology and Psychiatry*, *16*, 195–202.

Rose, M., Stedal, K., Reville, M.-C., van Noort, B. M., Kappel, V., Frampton, I., … Lask, B. (2016). Similarities and differences of neuropsychological profiles in children and adolescents with anorexia nervosa and healthy controls using cluster and discriminant function analyses. *Archives of Clinical Neuropsychology*, *31, 877-895.*

Rosenberg, H., McDonald, S., Dethier, M., Kessels, R. P., & Westbrook, R. F. (2014). Facial emotion recognition deficits following moderate–severe traumatic brain injury (TBI): Re-examining the valence effect and the role of emotion intensity. *Journal of the International Neuropsychological Society*, *20*, 994-1003.

Rothschild-Yakar, L., Eviatar, Z., Shamia, A., & Gur, E. (2011). Social cognition in eating disorders: encoding and representational processes in binging and purging patients. *European Eating Disorders Review*, *19*, 75-84.

Russell, T. A., Schmidt, U., Doherty, L., Young, V., & Tchanturia, K. (2009). Aspects of social cognition in anorexia nervosa: affective and cognitive theory of mind. *Psychiatry Research*, *168*, 181-185.

Sherratt, K. A. L., & MacLeod, A. K. (2013). Underlying motivation in the approach and avoidance goals of depressed and non-depressed individuals. *Cognition & Emotion*, *27*, 1432–1440.

Simmons, A. N., Fitzpatrick, S., Strigo, I. A., Potterat, E. G., Johnson, D. C., Matthews, S. C., ... & Paulus, M. P. (2012). Altered insula activation in anticipation of changing

emotional states: neural mechanisms underlying cognitive flexibility in special operations forces personnel. *Neuroreport, 23*, 234-239.

Simmons, A., Strigo, I. A., Matthews, S. C., Paulus, M. P., & Stein, M. B. (2009). Initial evidence of a failure to activate right anterior insula during affective set-shifting in PTSD. *Psychosomatic Medicine, 71,* 373.

Skinner, A. L., & Benton, C. P. (2010). Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychological Science*, *21*, 1248-1253.

Surcinelli, P., Codispoti, M., Montebarocci, O., Rossi, N., & Baldaro, B. (2006). Facial emotion recognition in trait anxiety. *Journal of Anxiety Disorders*, *20*, 110-117.

Surguladze, S. A., Young, A. W., Senior, C., Brébion, G., Travis, M. J., & Phillips, M. L. (2004). Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. *Neuropsychology*, *18*, 212.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*, 1092-1097.

Southgate, L., Tchanturia, K., & Treasure, J. (2005). Building a model of the aetiology of eating disorders by translating experimental neuroscience into clinical practice. *Journal of Mental Health*, *14*, 553-566.

Strother, E., Lemberg, R., Stanford, S. C., & Turberville, D. (2012). Eating disorders in men: underdiagnosed, undertreated, and misunderstood. *Eating Disorders*, *20*, 346-355.

Swanson, S. A., Crow, S. J., Le Grange, D., Swendsen, J., & Merikangas, K. R. (2011). Prevalence and correlates of eating disorders in adolescents: Results from the national comorbidity survey replication adolescent supplement. *Archives of General Psychiatry*, *68*, 714-723.

Tchanturia, K., Doris, E., Mountford, V., & Fleming, C. (2015). Cognitive Remediation and Emotion Skills Training (CREST) for anorexia nervosa in individual format: self-reported outcomes. *BioMed Central Psychiatry*, *15*, 53.

Tierney, S., & Fox, J. R. (2009). Chronic anorexia nervosa: A Delphi study to explore practitioners' views. International. *Journal of Eating Disorders, 42*, 62-67.

Torres, S., Guerra, M. P., Lencastre, L., Roma-Torres, A., Brandão, I., Queirós, C., & Vieira, F. (2011). Cognitive processing of emotions in anorexia nervosa. *European Eating Disorders Review, 19*, 100-111.

Treasure, J., & Schmidt, U. (2013). The cognitive-interpersonal maintenance model of anorexia nervosa revisited: a summary of the evidence for cognitive, socio-emotional and interpersonal predisposing and perpetuating factors. *Journal of Eating Disorders*, *1*, 13.

Uekermann, J., Abdel-Hamid, M., Lehmkaemper, C., Vollmoeller, W., & Daum, I. (2008). Perception of affective prosody in major depression: a link to executive functions? *Journal of the International Neuropsychological Society*, *14*, 552-561.

Van Autreve, S., De Baene, W., Baeken, C., van Heeringen, C., Vervaet, M., Autreve, S. Van, … Vervaet, M. (2013). Do restrictive and bingeing/purging subtypes of anorexia nervosa differ on central coherence and set shifting? *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *21*, 308–314.

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, *17*, 3-28.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron, 40*, 655-664.

Wittkampf, K., van Ravesteijn, H., Baas, K., van de Hoogen, H., Schene, A., Bindels, P., ... van Weert, H. (2009). The accuracy of Patient Health Questionnaire-9 in detecting depression and measuring depression severity in high-risk groups in primary care. *General Hospital Psychiatry, 31*, 451-459

Yoon, K. L., & Zinbarg, R. E. (2007). Threat is in the eye of the beholder: Social anxiety and the interpretation of ambiguous facial expressions. *Behaviour Research and Therapy*, *45*, 839-847.

Zakzanis, K. K., Campbell, Z., & Polsinelli, A. (2010). Quantitative evidence for distinct cognitive impairment in anorexia nervosa and bulimia nervosa. *Journal of Neuropsychology*, *4*, 89-106.

Appendices

Appendix A – NHS Ethical Approval

London - Queen Square Research Ethics Committee

HRA NRES Centre Manchester

Barlow House

3rd Floor

4 Minshull Street

Manchester

M1 3DZ

Study title: Emotion recognition, set-shifting and anorexia nervosa

REC reference: 17/LO/0908

Protocol number: 1617/20

IRAS project ID: 220842

The Proportionate Review Sub-committee of the London - Queen Square Research Ethics Committee reviewed the above application on 25 May 2017.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact hra.studyregistration@nhs.net outlining the reasons for your request. Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Ethical opinion

On behalf of the Committee, the sub-committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA Approval (England)/ NHS permission for research is available in the Integrated Research Application System, www.hra.nhs.uk or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion").

The Sub-Committee reviewed the application and no material ethical issues were raised.

Approved documents

The documents reviewed and approved were:

Document	*Version*	*Date*
Copies of advertisement materials for research participants [Poster and Leaflet]	1	27 March 2017
IRAS Application Form [IRAS_Form_11052017]		11 May 2017
IRAS Checklist XML [Checklist_11052017]		11 May 2017
Letter from sponsor [Sponsorship Letter to CI]		04 May 2017
Other [Somerset Local Collaborator Dr Dominic Hiles CV.docx]	1	08 May 2017
Other [Oxford and Wiltshire Local Collaborator Dr Mags Cariss CV.docx]	1	08 May 2017
Other [Avon and Wiltshire Local Collaborator Dr Sanni Norweg CV.docx]	1	08 May 2017
Other [Debrief Sheet]	1	18 April 2017
Other [Somerset R&D Letter of support]	1	01 December 2016
Other [Professional Indemnity/Insurance]	1	14 November 2016
Other [Public Liability/Insurance]	1	14 November 2016
Other [Application Clarification]		17 May 2017
Participant consent form	1	27 March 2017
Participant information sheet (PIS)	1	27 March 2017
Referee's report or other scientific critique report [Study Proposal Review by University of Exeter]	1	08 May 2017
Research protocol or project proposal [Research Protocol]	1	27 March 2017
Research protocol or project proposal [Risk Protocol]	1	13 March 2017

Summary CV for Chief Investigator (CI) [Chief Investigator CV]	1	15 February 2017
Summary CV for student [Chief Investigator CV]	1	08 May 2017
Summary CV for supervisor (student research) [Academic Supervisor Dr Nick Moberly CV]	1	08 May 2017
Summary CV for supervisor (student research) [Cornwall Local Collaborator and Academic Supervisor Dr Ian Frampton CV]	1	08 May 2017
Summary, synopsis or diagram (flowchart) of protocol in non technical language [Study Flow Chart]	1	07 May 2017
Validated questionnaire [BFNE-II]	1	07 May 2017
Validated questionnaire [PHQ9]	1	07 May 2017
Validated questionnaire [GAD7]	1	08 May 2017
Validated questionnaire [EDE-Q]	1	08 May 2017
Validated questionnaire [TAS-20]	1	08 May 2017

Membership of the Proportionate Review Sub-Committee

The members of the Sub-Committee who took part in the review are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document “After ethical review – guidance for researchers” gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

Notifying substantial amendments

Adding new sites and investigators

Notification of serious breaches of the protocol
Progress and safety reports

Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:

http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

With the Committee's best wishes for the success of this project.

17/LO/0908 **Please quote this number on all correspondence**

Yours sincerely

Signed on behalf of the Chair, Ms Danielle Wilson

Email: nrescommittee.london-queensquare@nhs.net

Enclosures:

List of names and professions of members who took part in the review "After ethical review – guidance for researchers" [SL-AR2]

Appendix B – University Ethical Approval

Application ID: **eCLESPsy000121 v4.1**
Title: **Emotion recognition, set-shifting and anorexia nervosa.**

Your e-Ethics application has been reviewed by the CLES Psychology Ethics Committee.

The outcome of the decision is: **Favourable**

Potential Outcomes

Favourable:	The application has been granted ethical approval by the Committee. The application will be flagged as Closed in the system. To view it again, please select the tick box: View completed
Favourable, with conditions:	The application has been granted ethical approval by the Committee under the provision of certain conditions. These conditions are detailed below.
Provisional:	You have **not** been granted ethical approval. The application needs to be amended in light of the Committee's comments and re-submitted for Ethical review.
Unfavourable:	You have **not** been granted ethical approval. The application has been **rejected** by the Committee. The application needs to be amended in light of the Committee's comments and resubmitted / or you need to complete a new application.

Please view your application here and respond to comments as required. You can download your outcome letter by clicking on the 'PDF' button on your eEthics Dashboard.

If you have any queries please contact the CLES Psychology Ethics Chair:
Lisa Leaver L.A.Leaver@exeter.ac.uk

Kind regards,
CLES Psychology Ethics Committee

Appendix C – Study Questionnaires

ASDP BFNE-II **1** of **1**

Initials/ID #:_____

Date:_____

Brief Fear of Negative Evaluation-II

(Carleton, Collimore, & Asmundson, 2007)

Please circle the number that best corresponds to how much you agree with each item.

	Not at all characteristic of me	A little characteristic of me	Somewhat characteristic of me	Very characteristic of me	Entirely characteristic of me
1. I worry about what other people will think of me even when I know it doesn't make any difference.	1	2	3	4	5
2. It bothers me when people form an unfavourable impression of me.	1	2	3	4	5
3. I am frequently afraid of other people noticing my shortcomings.	1	2	3	4	5
4. I worry about what kind of impression I make on people.	1	2	3	4	5
5. I am afraid that others will not approve of me.	1	2	3	4	5
6. I am afraid that other people will find fault with me.	1	2	3	4	5
7. I am concerned about other people's opinions of me.	1	2	3	4	5
8. When I am talking to someone, I worry about what they may be thinking about me.	1	2	3	4	5
9. I am usually worried about what kind of impression I make.	1	2	3	4	5
10. If I know someone is judging me, it tends to bother me.	1	2	3	4	5
11. Sometimes I think I am too concerned with what other people think of me.	1	2	3	4	5
12. I often worry that I will say or do wrong things.	1	2	3	4	5

Score:_____

EATING DISORDER EXAMINATION QUESTIONNAIRE (EDE-Q 6.0)

EATING QUESTIONNAIRE

Instructions: The following questions are concerned with the past four weeks (28 days) only. Please read each question carefully. Please answer all the questions. Thank you.

Questions 1 to 12: Please circle the appropriate number on the right. Remember that the questions only refer to the past four weeks (28 days) only.

	On how many of the past 28 days …..	**No days**	**1-5 days**	**6-12 days**	**13-15 days**	**16-22 days**	**23-27 days**	**Every day**
1	Have you been deliberately trying to limit the amount of food you eat to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6
2	Have you gone for long periods of time (8 waking hours or more) without eating anything at all in order to influence your shape or weight?	0	1	2	3	4	5	6
3	Have you tried to exclude from your diet any foods that you like in order to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6
4	Have you tried to follow definite rules regarding your eating (for example, a calorie limit) in order to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6
5	Have you had a definite desire to have an empty stomach with the aim of influencing your shape or weight?	0	1	2	3	4	5	6
6	Have you had a definite desire to have a totally flat stomach?	0	1	2	3	4	5	6
7	Has thinking about food, eating or calories made it very difficult to concentrate on things you are interested in (for example, working, following a conversation, or reading)?	0	1	2	3	4	5	6

8	Has thinking about <u>shape or weight</u> made it very difficult to concentrate on things you are interested in (for example, working, following a conversation, or reading)?	0	1	2	3	4	5	6
9	Have you had a definite fear of losing control over eating?	0	1	2	3	4	5	6
10	Have you had a definite fear that you might gain weight?	0	1	2	3	4	5	6
11	Have you felt fat?	0	1	2	3	4	5	6
12	Have you had a strong desire to lose weight?	0	1	2	3	4	5	6

Questions 13-18: Please fill in the appropriate number in the boxes on the right. Remember that the questions only refer to the past four weeks (28 days).

Over the past four weeks (28 days) ……

13	Over the past 28 days, how many <u>times</u> have you eaten what other people would regard as an <u>unusually large amount of food</u> (given the circumstances)?	**……** **……** **…..**
14	….. On how many of these times did you have a sense of having lost control over your eating (at the time that you were eating)?	**……** **……** **…..**
15	Over the past 28 days, on how many **<u>DAYS</u>** have such episodes of overeating occurred (i.e., you have eaten an unusually large amount of food <u>and</u> have had a sense of loss of control at the time)?	**……** **……** **…..**
16	Over the past 28 days, how many <u>times</u> have you made yourself sick (vomit) as a means of controlling your shape or weight?	**……** **……** **…..**

17 Over the past 28 days, how many <u>times</u> have you taken laxatives as a means of controlling your shape or weight?

……
……
…..

18 Over the past 28 days, how many <u>times</u> have you exercised in a "driven" or "compulsive" way as a means of controlling your weight, shape or amount of fat, or to burn off calories?

……
……
…..

Questions 19 to 21: Please circle the appropriate number. <u>Please note that for these questions the</u> <u>term "binge eating" means</u> eating what others would regard as an unusually large amount of food for the circumstances, accompanied by a sense of having lost control over eating.

19	Over the past 28 days, on how many days have you eaten in secret (ie, furtively)?	No days	1-5 days	6-12 days	13-15 days	16-22 days	23-27 days	Every day
	….. Do not count episodes of binge eating	0	1	2	3	4	5	6
20	On what proportion of the times that you have eaten have you felt guilty (felt that you've done wrong) because of its effect on your shape or weight?	None of the times	A few of the times	Less than half	Half of the times	More than half	Most of the time	Every time
	….. Do not count episodes of binge eating	0	1	2	3	4	5	6
21	Over the past 28 days, how concerned have you been about other people seeing you eat?	Not at all		Slightly	Moderately		Markedly	
	….. Do not count episodes of binge eating	0	1	2	3	4	5	6

Questions 22 to 28: Please circle the appropriate number on the right. Remember that the questions only refer to the past four weeks (28 days).

	Over the past 28 days …..	Not at all		Slightly		Moderately		Markedly
22	Has your weight influenced how you think about (judge) yourself as a person?	0	1	2	3	4	5	6
23	Has your shape influenced how you think about (judge) yourself as a person?	0	1	2	3	4	5	6
24	How much would it have upset you if you had been asked to weigh yourself once a week (no more, or less, often) for the next four weeks?	0	1	2	3	4	5	6
25	How dissatisfied have you been with your weight?	0	1	2	3	4	5	6
26	How dissatisfied have you been with your shape?	0	1	2	3	4	5	6
27	How uncomfortable have you felt seeing your body (for example, seeing your shape in the mirror, in a shop window reflection, while undressing or taking a bath or shower)?	0	1	2	3	4	5	6
28	How uncomfortable have you felt about others seeing your shape or figure (for example, in communal changing rooms, when swimming, or wearing tight clothes)?	0	1	2	3	4	5	6

What is your weight at present? (Please give your best estimate.) ……
What is your height? (Please give your best estimate.) …… If female: Over the past three-to-four months have you missed any menstrual periods? ……

If so, how many? …… Have you been taking the “pill”? ……

THANK YOU

Generalized Anxiety Disorder 7-item (GAD-7) scale

Over the last 2 weeks, how often have you been bothered by the following problems?	Not at all sure	Several days	Over half the days	Nearly every day
1. Feeling nervous, anxious, or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it's hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid as if something awful might happen	0	1	2	3
Add the score for each column	+	+	+	
Total Score *(add your column scores)* =				

If you checked off any problems, how difficult have these made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all _____

Somewhat difficult _____

Very difficult _____

Extremely difficult _____

The Patient Health Questionnaire (PHQ-9)

Over the past 2 weeks, how often have you been bothered by any of the following problems?	Not At all	Several Days	More Than Half the Days	Nearly Every Day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed or hopeless	0	1	2	3
3. Trouble falling asleep, staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself - or that you're a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or, the opposite - being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3
Column Totals	_____	+ _____	+ _____	
Add Totals Together	_____			

TORONTO ALEXITYMIA SCALE (TAS-20)

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neither Disagree or Agree**	**Agree**	**Strongly Agree**

_____ 1. I am often confused about what emotion I am feeling.
_____ 2. It is difficult for me to find the right words for my feelings.
_____ 3. I have physical sensations that even doctors don't understand.
_____ 4. I am able to describe my feelings easily
_____ 5. I prefer to analyze problems rather than just describe them.
_____ 6. When I am upset, I don't know if I am sad, frightened, or angry
_____ 7. I find it hard to describe how I feel about people.
_____ 8. I prefer to just let things happen rather than to understand why they turned out that way.
_____ 9. I have feelings that I can't quite identify.
_____ 10. Being in touch with emotions is essential.
_____ 11. I am often puzzled by sensations in my body.
_____ 12. People tell me to describe my feelings more.
_____ 13. I don't know what's going on inside me.
_____ 14. I often don't know why I am angry.
_____ 15. I prefer talking to people about their daily activities rather than their feelings.
_____ 16. I prefer to watch "light" entertainment shows rather than psychological dramas.
_____ 17. It is difficult for me to reveal my innermost feelings, even to close friends.
_____ 18. I can feel close to someone, even in moments of silence.
_____ 19. I find examination of my feelings useful in solving personal problems
_____ 20. Looking for hidden meanings in movies or plays distracts from their enjoyment.

Study Information Sheet

Study Title:

Emotion recognition, set-shifting and anorexia nervosa

Invitation:

You are invited to participate in a study that aims to explore how people recognise emotions in other people's faces and how this may relate to eating behaviour and other thinking skills.

The study is being conducted by Royston Hall, Trainee Clinical Psychologist as part of a PhD thesis, supervised by Dr Ian Frampton and Dr Nick Moberly at the University of Exeter.

Before you decide whether or not you wish to participate in this study, it is important for you to understand why the research is being done and what it will involve. Please take the time to read the following information carefully and discuss it with others if you wish.

Participants can choose to enter a prize draw for one of two £50 Amazon vouchers.

'What is the purpose of this study?

The purpose of this study is to explore whether there is a relationship between emotion recognition ability, thinking skills and eating behaviours. There is already some evidence that there are differences in how people with anorexia recognise and experience emotions in other people's faces but it is unclear whether this is related to any other kinds of thinking skills. Understanding these relationships may help researchers inform, develop and enhance clinical treatment for individuals with anorexia nervosa.

'What does this study involve?'

If you agree to participate in this study, you will be asked to sign the attached Consent Form. A session will be booked at a time and location most suitable to you such as your local NHS service site or University.

The study involves completing some short computer tasks and brief questionnaires. The computer tasks are simple tasks that will ask you to respond to different images in different ways. These tasks take between 30-45 minutes. The questionnaires will ask questions about your thoughts, feelings and behaviours in different situations. The questionnaires typically take 10-20 minutes. Overall the study will take 45-65 minutes to complete and you are only required to participate at one time.

The session will be run by a trained researcher who will be present at all times to talk through the task and answer any questions you may have.

'Are there risks to me in taking part in this study?'

The tasks are brief and simple and should not cause significant distress. However, if you do feel distressed at any point during the tasks you are free to stop and withdraw at any time.

Some people may find the clinical questionnaires concerning, especially if the questions raise worries about your own health. Information on who to contact if you are concerned will be provided in a debrief sheet and you are welcome to discuss your concerns with the researcher who can also advise you.

All information will remain anonymous and confidential. You are free to withdraw your information at any time. Taking part in the study will not affect your treatment or your future access to treatment in any way.

'Will I benefit from the study?'

There are no expected direct benefits for participants of this research. However, it will help improve our understanding of the complex origin of AN.

If you would like to know the outcome of the study, tick the box on the consent form and provide your email address and you will be updated upon completion of the study.

By participating you are able to be entered into the prize draw for £50 worth of Amazon vouchers. If you would like to be entered into the draw, please tick the box and fill in your e-mail address on the Consent Form. Your email address will be kept in a password protected excel sheet, it will not be shared with any other researcher or organisation and will be deleted once a winner has been selected.

'What if I don't want to take part in this study?'

Participation in this study is voluntary. It is completely up to you whether or not you participate. Your decision not to participate will be respected and will not affect your current or future access to healthcare.

'What if I participate and want to withdraw later?'

You are free to withdraw from the study anytime without consequence. Note, however, that because data will be collected and stored in de-identified form, any data that you have provided will not be able to be withdrawn once it has reached the stage of analysis.

'How will my confidentiality be protected?'

You will not be identified by name and all data collected will be stored in a password protected file on any laptop or an encrypted portable external drive. It will be downloaded on password protected files on the University of Exeter network which is itself password protected. The electronic data files will be deleted after 5 years from the University of Exeter networked storage. All paper files will be converted to electronic databases and shredded immediately.

'What happens with the results?'

The results will be used to help write a thesis as part of the DClinPsy requirements for a Doctorate in Clinical Psychology. If appropriate, the results may be published in a peer reviewed journal such as the European Eating Disorders Review. In all cases, information will be provided in such a way that you cannot be identified.

Aggregated results of the study will be provided to you, if you wish.

'What should I do if I want to discuss this study further before I decide?'

If you would like to know more about the study at any stage, please do not hesitate to contact Royston Hall on rh487@exeter.ac.uk or 07908383786.

Complaints

If you have any complaints about the way in which this study has been carried out please contact Nick Moberly at N.J.Moberly@exeter.ac.uk, or the University of Exeter's Research Ethics and Governance Manager, Gail Seymour: Email: G.M.Seymour@exeter.ac.uk; Tel: 01392 726621

Who is organising, funding and ethically reviewing the project?

The study is sponsored by The University of Exeter. It is funded by the Economic and Social Research Council (ESRC).

All research involving NHS patients is looked at by an independent Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given a favourable opinion by Ms Danielle Wilson of Queen Square NHS Ethics Committee, London.

If you would like any independent advice about participating in research you can contact PALS (the local Patient Advice and Liaison Service), or INVOLVE at www.invo.org.uk/

Appendix E – Participant Consent Form

UNIVERSITY OF EXETER

IRAS ID: 220842 Centre Number: Study Number:

Participant Identification Number for this trial:

CONSENT FORM

Title of Project: Emotion recognition, set-shifting and anorexia nervosa.

1. I confirm that I have read the information sheet dated 27.03.2017 (version 1) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that I can ask for my data to be destroyed at any point without penalty. ☐
4. I understand that my participation in this study will be anonymised and that I will not be able to be identified in any way to anyone outside the research team or in any future publications. ☐
5. I understand that my participation is confidential unless there is a risk of harm to myself or others. ☐
6. I agree to take part in the above study. ☐

_____	_____	_____
Name of Participant	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature

I would like to enter the prize draw for one of two £50 Amazon Vouchers. ☐

I would like to be informed of the outcome of this study. ☐

If you answered yes to either of these questions, please provide your email address below:

_____

Appendix F – Dissemination Statement

Results will be disseminated on multiple levels. First, following data collection and analysis, participants who expressed interest and consented to be contacted will be provided with a letter describing in lay terms the main findings of the present study.

Secondly, the findings of the present study will be presented at the local services that served as recruitment sites. The present research will also be presented to the Exeter Brain's meeting of Anorexia Research. A talk at Exeter University has also been scheduled for June 2018 to present the findings to colleagues and other professionals.

Finally, revised versions of the final literature review and empirical paper will be submitted for publication to the peer reviewed journal *European Eating Disorders Review.*

Appendix G – Preparation and Submission Requirements for the European Eating Disorders Review

1. SUBMISSION

Authors should kindly note that submission implies that the content has not been published or submitted for publication elsewhere except as a brief abstract in the proceedings of a scientific meeting or symposium.

Once the submission materials have been prepared in accordance with the Author Guidelines, manuscripts should be submitted online at http://mc.manuscriptcentral.com/erv

2. AIMS AND SCOPE

European Eating Disorders Review provides an international forum for disseminating cutting-edge theoretical and empirical research that significantly advances understanding of the relationship between Eating Disorders and Abnormal Eating/Weight conditions and well-being in humans.

European Eating Disorders Review publishes authoritative and accessible articles, from all over the world, which review or report original research that has implications for the treatment and care of people with eating disorders and obesity, and articles which report innovations and experience in the clinical management of eating disorders. The journal focuses on implications for best practice in diagnosis and treatment. The journal also provides a forum for discussion of the causes and prevention of eating disorders, and related health policy.

Authors may submit original theoretical systematic reviews, methodological, or empirical research articles (7000 words or less) or short communications (3000 words or less). The journal also publishes invited conceptual reviews from leading worldwide researchers in the field of Eating Disorders and/or Obesity. The aims of the journal are to offer a channel of communication between researchers, practitioners, administrators and policymakers who need to report and understand developments in the field of eating disorders.

The journal

- Reports on useful research and experience related to the treatment and prevention of eating disorders in primary care and hospital settings, with special attention to therapy oriented translational research, high quality reviews, clinical trials and pilot innovative therapy approaches.
- Provides information about 'good practice' and systematic reviews.
- Offers a forum for new thinking about the nature, incidence, diagnosis and clinical management of eating disorders (namely anorexia nervosa, bulimia nervosa, binge eating disorders, OSFED and other abnormal eating or feeding behaviors associated with childhood and obesity.

3. MANUSCRIPT CATEGORIES AND REQUIREMENTS

Research articles reporting new research of relevance as set out in the aims and scope should not normally exceed 6000 words (excluding abstract, references, tables or figures), with no more than five tables or illustrations. They should conform to the conventional layout: title page, Abstract, Introduction and Aims, Method, Results, Discussion, Acknowledgements and References. Each of these elements should start on a new page.
Word Limit: 6,000 (excluding abstract, references, tables or figures).

Abstract: 200 words.
References: up to 60.

Review articles: Systematic and meta-analytic review papers are welcomed if they critically review the available literature in a topic than will enhance clinical practice. Articles should have clear focus and enough number of studies should be available for a substantive review paper. Studies that only describe or list previous studies without a critical overview of the literature will not be considered.
Word Limit: 5,000 (excluding abstract, references, tables or figures).
Abstract: 200 words.
References: up to 100.
Figures/Tables: 5 maximum, but should be appropriate to the material covered. Additional tables might be included as supplementary information, if needed. Review articles must follow the PRISMAGuidelines. Authors may want to have a look at the review check lists that reviewers when assessing review articles.

Brief reports should concisely present the essential findings of the author's work and be compromised of the following sections: Abstract, Introduction and Aims, Method, Results, Discussion, and References. Tables and/or figures should be kept to a minimum, in number and size, and only deal with key findings. In some cases authors may be asked to prepare a version of the manuscript with extra material to be included in the online version of the review (as supplementary files). Submissions in this category should not normally exceed 2500 words in length.

Brief reports bring with them a whole host of benefits including: quick and easy submission, administration centralised and reduced and significant decrease in peer review times, first publication priority (this type of manuscript will be published in the next available issue of the journal).

Case Reports The journal does not accept case reports for publication. Authors of case reports are encouraged to submit to the Wiley Open Access journal, Clinical Case Reports www.clinicalcasesjournal.com which aims to directly improve health outcomes by identifying and disseminating examples of best clinical practice.

4. PREPARING THE SUBMISSION

Cover Letters

Cover letters are not mandatory; however, they may be supplied at the author’s discretion.

Parts of the Manuscript

The manuscript should be submitted in separate files: title page; main text file; figures.

Main Text File

The text file should be presented in the following order:

i. A short title that contains the major key words. The title should not contain abbreviations (see Wiley's best practice SEO tips);
ii. A short running title of less than 40 characters;
iii. The full names of the authors;
iv. The author's institutional affiliations where the work was conducted, with a footnote for the author’s present address if different from where the work was conducted;
v. The corresponding author’s contact email address and telephone number;
vi. Acknowledgments;
vii. Conflict of Interest statement (for all authors)
viii. Names and grant numbers of any sources of funding or support in the form of grants, equipment, drugs etc.

Authorship

Please refer to the journal’s authorship policy the Editorial Policies and Ethical Considerations section for details on eligibility for author listing eligibility.

Acknowledgments

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section. Financial and material support should also be mentioned. Thanks to anonymous reviewers are not appropriate.

Conflict of Interest Statement

Authors will be asked to provide a conflict of interest statement during the submission process. For details on what to include in this section, see the section 'Conflict of Interest' in the Editorial Policies and Ethical Considerations section below. Submitting authors should ensure they liaise with all co-authors to confirm agreement with the final statement.

Main Text File

As papers are double-blind peer reviewed, the main text file should not include any information that might identify the authors.

The main text file should be presented in the following order:

i. Title, abstract, highlights and key words;
ii. Main text;
iii. References;
iv. Tables (each table complete with title and footnotes);
v. Figure legends;
vi. Appendices (if relevant).

Figures and supporting information should be supplied as separate files.

Abstract

All manuscripts should contain an abstract of up to 200 words. An **abstract** is a concise summary of the whole paper, not just the conclusions, and is understandable without reference to the rest of the paper. It should contain no citation to other published work. It must be structured, under the sub-headings: Objective; Method; Results; Conclusions.

Highlights

Highlights are mandatory for European Eating Disorders Review. These should appear as three bullet points that convey the core findings of the article.

Keywords

Include up to five **keywords** that describe your paper for indexing purposes.

References

References should be prepared according to the Publication Manual of the American Psychological Association (6th edition). This means in text citations should follow the author-date method whereby the author's last name and the year of publication for the source should appear in the text, for example, (Jones, 1998). The complete reference list should appear alphabetically by name at the end of the paper. A sample of the most common entries in reference lists appears below. Please note that a DOI should be provided for all references where available. For more information about APA referencing style, please refer to the APA FAQ. Please note that for journal articles, issue numbers are not included unless each issue in the volume begins with page one.

Tables

Tables should be self-contained and complement, not duplicate, information contained in the text. They should be supplied as editable files, not pasted as images. Legends should be concise but comprehensive – the table, legend, and footnotes must be understandable without reference to the text. All abbreviations must be defined in footnotes. Footnote symbols: †, ‡, §, ¶, should be used (in that order) and *, **, *** should be reserved for P-values. Statistical measures such as SD or SEM should be identified in the headings.

Figure Legends

Legends should be concise but comprehensive – the figure and its legend must be understandable without reference to the text. Include definitions of any symbols used and define/explain all abbreviations and units of measurement.

Figures

Although authors are encouraged to send the highest-quality figures possible, for peer-review purposes, a wide variety of formats, sizes, and resolutions are accepted. Click here for the basic figure requirements for figures submitted with manuscripts for initial peer review, as well as the more detailed post-acceptance figure requirements.

Additional Files

Appendices

Appendices will be published after the references. For submission they should be supplied as separate files but referred to in the text.

Supporting Information

Supporting information is information that is not essential to the article, but provides greater depth and background. It is hosted online and appears without editing or typesetting. It may include tables, figures, videos, datasets, etc. Click here for Wiley's FAQs on supporting information.

Note: if data, scripts, or other artefacts used to generate the analyses presented in the paper are available via a publicly available data repository, authors should include a reference to the location of the material within their paper.

If a manuscript describes a new approach and/or technological approach, authors are encouraged to include a small demo video – no more than 60 seconds long.

General Style Points

The following points provide general advice on formatting and style.

- **Language:** The language of the journal is English.
- **Abbreviations:** In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Initially, use the word in full, followed by the abbreviation in parentheses. Thereafter use the abbreviation only. Distinction should be made between capital and lower case letters, between the letter O and zero, between the letter I and number one and prime, between k and kappa.
- **Units of measurement:** Measurements should be given in SI or SI-derived units. Visit the Bureau International des Poids et Mesures (BIPM) website at www.bipm.fr for more information about SI units.
- **Numbers:** numbers under 10 are spelt out, except for: measurements with a unit (8mmol/l); age (6 weeks old), or lists with other numbers (11 dogs, 9 cats, 4 gerbils).
- **Trade Names:** Chemical substances should be referred to by the generic name only. Trade names should not be used. Drugs should be referred to by their generic names. If proprietary drugs have been used in the study, refer to these by their generic name, mentioning the proprietary name and the name and location of the manufacturer in parentheses.

Wiley Author Resources

Manuscript Preparation Tips: Wiley has a range of resources for authors preparing manuscripts for submission available here. In particular, authors may benefit from referring to Wiley's best practice tips on Writing for Search Engine Optimization.

Editing, Translation, and Formatting Support: Wiley Editing Services can greatly improve the chances of a manuscript being accepted. Offering expert help in English language editing, translation, manuscript formatting, and figure preparation, Wiley Editing Services ensures that the manuscript is ready for submission.

5. EDITORIAL POLICIES AND ETHICAL CONSIDERATIONS

Editorial Review and Acceptance

The acceptance criteria for all papers are the quality and originality of the research and its significance to journal readership. Manuscripts are double-blind peer reviewed. Papers will only be sent to review if the Editor-in-Chief determines that the paper meets the appropriate quality and relevance requirements.

Wiley's policy on the confidentiality of the review process is available here.

Decision Appeals

Any appeal against a decision on a manuscript should be filed by the corresponding author within 28 days of notification of the decision. The appeal should be in the form of a letter addressed to the Editor-in-Chief and submitted by email to the editorial office (EEDRedoffice@wiley.com). The letter should include clear and concise grounds for the appeal, including specific points of disagreement with the decision. The appeal will then be assessed by the editorial team, led by the Editor-in-Chief, and informed by the reviewer assessments and recommendation of the Associate Editors, where appropriate. Authors lodging an appeal will be informed of its outcome within 28 days. The decision will be final.

Data Storage and Documentation

European Eating Disorders Review encourages authors to share the data and other artefacts supporting the results in the paper by archiving it in an appropriate public repository. Authors should include a data accessibility statement, including a link to the repository they have used, in order that this statement can be published alongside their paper.

Authors can consult the global registry of research data repositories re3data.org to help them identify registered and certified repositories relevant to their subject areas.

Data Citation In recognition of the significance of data as an output of research effort, Wiley has endorsed the FORCE11 Data Citation Principles and is implementing a mandatory data citation policy. Journal policies should require data to be cited in the same way as article, book, and web citations and authors are required to include data citations as part of their reference list. Data citation is appropriate for data held within institutional, subject focused, or more general data repositories. It is not intended to take the place of community standards such as in-line citation of GenBank accession codes. When citing or making claims based on data, authors must refer to the data at the relevant place in the manuscript text and in addition provide a formal citation in the reference list. We recommend the format proposed by the Joint Declaration of Data Citation Principles

Authors; Year; Dataset title; Data repository or archive; Version (if any); Persistent identifier (e.g. DOI)

Human Studies and Subjects

For manuscripts reporting medical studies that involve human participants, a statement identifying the ethics committee that approved the study and confirmation that the study conforms to recognized standards is required, for example: Declaration of Helsinki; US Federal Policy for the Protection of Human Subjects; or European Medicines Agency Guidelines for Good Clinical Practice. It should also state clearly in the text that all persons gave their informed consent prior to their inclusion in the study.

Patient anonymity should be preserved. Photographs need to be cropped sufficiently to prevent human subjects being recognized (or an eye bar should be used). Images and information from individual participants will only be published where the authors have obtained the individual's free prior informed consent. Authors do not need to provide a copy of the consent form to the publisher; however, in signing the author license to publish, authors are required to confirm that consent has been obtained. Wiley has a standard patient consent form available for use.

Clinical Trial Registration

The journal requires that clinical trials are prospectively registered in a publicly accessible database and clinical trial registration numbers should be included in all papers that report their results. Authors are asked to include the name of the trial register and the clinical trial registration number at the end of the abstract. If the trial is not registered, or was registered retrospectively, the reasons for this should be explained.

Research Reporting Guidelines

Accurate and complete reporting enables readers to fully appraise research, replicate it, and use it. Authors are expected to adhere to the following research reporting standards.

- CONSORT checklist for reports of randomised trials and cluster randomised trials
- TREND checklist for non-randomised controlled trials
- PRISMA checklist for systematic reviews and meta-analyses • STROBE checklist for observational research
- COREQ checklist for qualitative studies
- SQUIRE checklist for quality improvement

See the EQUATOR Network for other study types.

Conflict of Interest

The journal requires that all authors disclose any potential sources of conflict of interest. Any interest or relationship, financial or otherwise that might be perceived as influencing an author's objectivity is considered a potential source of conflict of interest. These must be disclosed when directly relevant or directly related to the work that the authors describe in their manuscript. Potential sources of conflict of interest include, but are not limited to: patent or stock ownership, membership of a company board of directors, membership of an advisory board or committee for a company, and consultancy for or receipt of speaker's fees from a company. The existence of a conflict of interest does not preclude publication. If the authors have no conflict of interest to declare, they must also state this at submission. It is the responsibility of the corresponding author to review this policy with all authors and collectively to disclose with the submission ALL pertinent commercial and other relationships.

Funding

Authors should list all funding sources in the Acknowledgments section. Authors are responsible for the accuracy of their funder designation. If in doubt, please check the Open Funder Registry for the correct nomenclature: https://www.crossref.org/services/funder-registry/

The list of authors should accurately illustrate who contributed to the work and how. All those listed as authors should qualify for authorship according to the following criteria:

1. Have made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; and
2. Been involved in drafting the manuscript or revising it critically for important intellectual content; and
3. Given final approval of the version to be published. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content; and
4. Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section (for example, to recognize contributions from people who provided technical help, collation of data, writing assistance, acquisition of funding, or a department chairperson who provided general support). Prior to submitting the article all authors should agree on the order in which their names will be listed in the manuscript.

Additional Authorship Options:

In the case of joint or shared first authorship, a footnote should be added to the author listing, e.g. ‘X and Y should be considered joint first author’ or ‘X and Y should be considered joint senior author.’

ORCID

As part of the journal’s commitment to supporting authors at every step of the publishing process, the journal requires the submitting author (only) to provide an ORCID iD when submitting a manuscript. This takes around 2 minutes to complete. Find more information here.

Publication Ethics

This journal is a member of the Committee on Publication Ethics (COPE). Note this journal uses iThenticate's CrossCheck software to detect instances of overlapping and similar text in submitted manuscripts. Read Wiley'sTop 10 Publishing Ethics Tips for Authors here. Wiley's Publication Ethics Guidelines can be found here.

6. AUTHOR LICENSING

If your paper is accepted, the author identified as the formal corresponding author will receive an email prompting them to log in to Author Services, where via the Wiley Author Licensing Service (WALS) they will be required to complete a copyright license agreement on behalf of all authors of the paper.

Authors may choose to publish under the terms of the journal's standard copyright agreement, or OnlineOpen under the terms of a Creative Commons License.

General information regarding licensing and copyright is available here. To review the Creative Commons License options offered under OnlineOpen, please click here. (Note that certain funders mandate that a particular type of CC license has to be used; to check this please click here.)

Self-Archiving definitions and policies. Note that the journal's standard copyright agreement allows for self-archiving of different versions of the article under specific conditions. Please click here for more detailed information about self-archiving definitions and policies.

Open Access fees: If you choose to publish using OnlineOpen you will be charged a fee. A list of Article Publication Charges for Wiley journals is available here.

Funder Open Access: Please click here for more information on Wiley's compliance with specific Funder Open Access Policies.

7. PUBLICATION PROCESS AFTER ACCEPTANCE

Accepted article received in production

When an accepted article is received by Wiley's production team, the corresponding author will receive an email asking them to login or register with Wiley Author Services. The author will be asked to sign a publication license at this point.

Proofs

Once the paper is typeset, the author will receive an email notification with the URL to download a PDF typeset page proof, as well as associated forms and full instructions on how to correct and return the file.

Please note that the author is responsible for all statements made in their work, including changes made during the editorial process – authors should check proofs carefully. Note that proofs should be returned within 48 hours from receipt of first proof.

Publication Charges

Colour figures. Colour figures may be reproduced in colour online free of charge. Please note, however, that it is preferable that line figures (e.g. graphs and charts) are supplied in black and white so that they are legible if printed by a reader in black and white. If an author would prefer to have figures printed in colour in hard copies of the journal, a fee will be charged by the Publisher.

Early View

The journal offers rapid speed to publication via Wiley's Early View service. Early View (Online Version of Record) articles are published on Wiley Online Library before inclusion in an issue. Note there may be a delay after corrections are received before the article appears online, as Editors also need to review proofs. Once the article is published on Early View, no further changes to the article are possible. The Early View article is fully citable and carries an online publication date and DOI for citations.

8. POST PUBLICATION

Access and sharing

When the article is published online:

• The author receives an email alert (if requested).
• The link to the published article can be shared through social media.
• The author will have free access to the paper (after accepting the Terms & Conditions of use, they can view the article).
• The corresponding author and co-authors can nominate up to ten colleagues to receive a publication alert and free online access to the article.

Promoting the Article

To find out how to best promote an article, click here.

Measuring the Impact of an Article

Wiley also helps authors measure the impact of their research through specialist partnerships with Kudos and Altmetric.

9. EDITORIAL OFFICE CONTACT DETAILS

EEDRedoffice@wiley.com